# WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



# INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

WO 00/18905 (51) International Patent Classification  $^7$ : (11) International Publication Number: C12N 15/10, 15/25, 15/62, C07K 14/535, A1 6 April 2000 (06.04.00) (43) International Publication Date: 14/005

US

- PCT/US99/20891 (21) International Application Number:
- 24 September 1999 (24.09.99) (22) International Filing Date:
- 60/101,908

25 September 1998 (25.09.98)

- (71) Applicant (for all designated States except US): G.D. SEARLE & CO. [US/US]; Corporate Patent Department, P.O. Box 5110, Chicago, IL 60680-5110 (US).
- (72) Inventor; and

(30) Priority Data:

- (75) Inventor/Applicant (for US only): LEE, Stephen, C. [US/US]; 828 Fernview Drive, Creve Coeur, MO 63141 (US).
- (74) Agents: BAUER, S., Christopher et al.; G.D. Searle & Co., Corporate Patent Dept., P.O. Box 5110, Chicago, IL 60680-5110 (US).
- (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

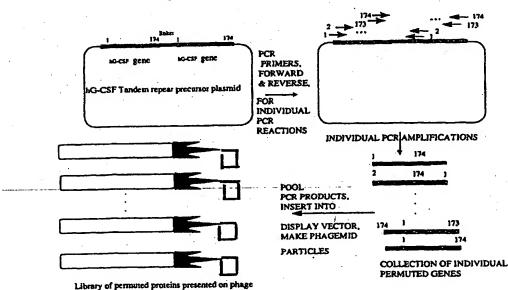
#### Published

With international search report. Before the expiration of the time limit for amending the

claims and to be republished in the event of the receipt of amendments.

## (54) Title: METHOD OF PRODUCING PERMUTEINS BY SCANNING PERMUTAGENESIS

## Constructing a scanning permutagenesis display library



#### (57) Abstract

A method of producing circularly-permuted proteins (permuteins) by scanning permutagenesis comprises making and inserting a series of circularly-permuted genes into a display vector, expressing these genes such that the gene products are localized to the surface of the display vector, generating a library of display vectors presenting the permuted protein, affinity-selecting the display vectors with a target protein that can bind the permuted protein, isolating and analyzing clones of selected display vectors to identify the circularly-permuted protein. The invention further discloses methods of expressing and uses of permuteins.

BEST AVAILABLE COP

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB .	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG .	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	. TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT .	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL.	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	· UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
,cc	Congo	KE	Кепуа	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP -	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		•
CN	China	KR	Republic of Korea	PT .	Portugal		•
CU	Cuba	KZ	Kazakstan	RO	Romania		
cz	Czech Republic	LC	Saint Lucia	RU '	Russian Federation		
DE	Germany	L	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EB	Estonia	LR	Liberia	SG	Singapore		
					•		

WO 00/18905 PCT/US99/20891

# Method of producing permuteins by scanning permutagenesis

#### Priority

The present application claims priority under Title 35, United States Code, § 119 of United States Provisional Application Serial No. 60/101,908, filed September 25, 1998.

### Field of the invention

A method of producing circularly-permuted proteins (permuteins) by scanning permutagenesis comprises making and inserting a series of circularly-permuted genes into a display vector, expressing these genes such that the gene products are localized to the surface of the display vector, generating a library of display vectors presenting the permuted protein, affinity-selecting the display vectors with a target protein that can bind the permuted protein, isolating and analyzing clones of selected display vectors to identify the circularly-permuted protein. The invention further discloses methods of expressing and uses of permuteins.

### Background of the invention

#### Protein permutagenesis

Circularly permuted proteins are made by reordering the primary sequence of a parent protein. The amino and carboxy terminal ends of the parent protein are joined by a peptide linker and new amino and carboxy terminal ends are generated at other positions in the sequence. This technique of generating variants has been applied to a wide variety of proteins (Table 1).

Circularly permuted proteins, in many cases, are structurally and functionally similar to their non-permuted parent molecule after they undergo refolding. The information necessary to direct the folding of proteins into tertiary structures is present in secondary structural domains. Vectorial folding of proteins from their native amino to carboxy ends is not often observed. The ability of permuteins to retain structural and functional properties is remarkable, extending earlier observations on the plasticity of proteins with respect to amino acid

30

25

5

10

15

20

10

15

20

25

30

35

substitutions (Olins P.O. et al., *J. Biol. Chem.* 270: 23754-23760, 1995; Lowman and Wells, *J. Mol. Biol.* 234: 564-578, 1993) and short amino acid insertions (Sondek, J. and D. Shortle, *Proteins* 7: 387-393, 1990; Shortle, D. and J. Sondek, *Curr. Opin. Biotechnol.* 6: 299-305).

## Protein sequence reorganization

Rearrangements of DNA sequences serve an important role in evolution by generating a diversity of new proteins differing in structure and function. Gene duplication and exon shuffling, for example, generate diversity and provide organisms with a competitive advantage since the basal mutation rate is low (Doolittle, *Protein Science* 1: 191-200, 1992).

Recombinant DNA methods have facilitated studies on the effect of sequence transposition on protein folding, structure, and function. rearrangement of proteins using this approach was described by Goldenberg and Creighton (J. Mol. Biol. 165:407-413, 1983). A new N-terminus is selected at an internal site (breakpoint) of the original sequence, the new sequence having the same order of amino acids as the original from the breakpoint until it reaches an amino acid that is at or near the original C-terminus. At this point the new sequence is joined, either directly or through an additional portion of sequence (linker), to an amino acid that is at or near the original N-terminus, and the new sequence continues with the same sequence as the original until it reaches a point that is at or near the amino acid that was N-terminal to the breakpoint site of the original sequence, this residue forming the new C-terminus of the chain. Similar approaches have also been used in other studies (Cunningham et al., Proc. Natl. Acad. Sci. U.S.A. 76:3218-3222, 1979; Teather & Erfle, J. Bacteriol. 172: 3837-3841, 1990; Schimming et al., Eur. J. Biochem. 204: 13-19, 1992; Yamiuchi and Minamikawa, FEBS Lett. 260:127-130, 1991: MacGregor et al., FEBS Lett. **378**:263-266, 1996).

These general approaches have been applied to proteins which range in size from 58 to 462 amino acids (Goldenberg & Creighton, J. Mol. Biol. 165:407-413, 1983; Li & Coffino, Mol. Cell. Biol. 13:2377-2383, 1993). The proteins represent a broad range of structural classes, including proteins that contain predominantly alpha helix (interleukin-4; Kreitman et al., Cytokine 7:311-318, 1995), beta sheet (interleukin-1; Horlick et al., Protein Eng. 5:427-431, 1992), or mixtures of the two types of secondary structures (yeast phosphoribosyl anthranilate isomerase; Luger et al., Science 243:206-210, 1989).

WO 00/18905 PCT/US99/20891

3

5

10

15

20

25

30

35

Although broad categories of protein function are represented in these sequence reorganization studies, the results of these studies have been highly variable. In many cases substantially lower activity, solubility, or thermodynamic dihydrofolate reductase, (E. coli observed stability were transcarbamoylase, phosphoribosyl anthranilate isomerase, glyceraldehyde-3phosphate dehydrogenase, ornithine decarboxylase, ompA, yeast phosphoglycerate dehydrogenase). In other cases, the sequence rearranged protein appeared to have many nearly identical properties as its natural counterpart (basic pancreatic trypsin inhibitor, T4 lysozyme, ribonuclease T1, Bacillus β-glucanase, interleukin-1β, α-spectrin SH3 domain, pepsinogen, interleukin-4). In exceptional cases, an unexpected improvement over some properties of the natural sequence was observed, e.g., the solubility and refolding rate for rearranged a-spectrin SH3 domain sequences, and the receptor affinity and anti-tumor activity of transposed interleukin-4-Pseudomonas exotoxin fusion molecule (Kreitman et al., Proc. Natl. Acad. Sci. U.S.A. 91:6889-6893, 1994; Kreitman et al., Cancer Res. 55:3357-3363, 1995).

The primary motivation for reorganization studies has been to study the role of short-range and long-range interactions in protein folding and stability. Sequence rearrangements of this type convert a subset of interactions that are long-range in the original sequence into short-range interactions in the new sequence, and vice versa. The fact that many of these sequence rearrangements are able to attain a conformation with at least some activity is persuasive evidence that protein folding occurs by multiple folding pathways (Viguera et al., J. Mol. Biol. 247:670-681, 1995). In the case of the SH3 domain of alpha-spectrin, choosing new termini at locations that corresponded to beta hairpin turns resulted in proteins with slightly less stability, but which were nevertheless able to fold.

The positions of the internal breakpoints used in the studies cited above are found exclusively on the surface of proteins, and are distributed throughout the linear sequence without any obvious bias towards the ends or the middle (the variation in the relative distance from the original N-terminus to the breakpoint is ca. 10 to 80% of the total sequence length). The linkers connecting the original N-and C-termini in these studies have ranged from 0 to 9 residues. In one case (Yang & Schachman, Proc. Natl. Acad. Sci. U.S.A. 90:11980-11984, 1993), a portion of sequence has been deleted from the original C-terminal segment, and the connection made from the truncated C-terminus to the original N-terminus. Flexible hydrophilic residues such as Gly and Ser are frequently used in the linkers. Viguera et al.(J. Mol. Biol. 247:670-681, 1995) compared joining the

10

15

20

25

30

original N- and C- termini with 3- or 4-residue linkers; the 3-residue linker was less thermodynamically stable. Protasova et al. (*Protein Eng.* 7:1373-1377, 1994) used 3- or 5-residue linkers in connecting the original N-termini of *E. coli* dihydrofolate reductase; only the 3-residue linker produced protein in good yield.

Protein permutagenesis can be used to optimize the activity of fusion proteins or proteins conjugated to other molecules. A fusion between interleukin-4 — (IL-4) and *Pseudomonas* exotoxin has been permuted resulting in a protein that has the first amino acid of the IL-4 domain at position 38 and the new carboxy end occurs at amino acid position 37 (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994). The permuted fusion has increased affinity for the IL-4 receptor, increased cytotoxicity to IL-4 receptor bearing renal carcinoma cells, and increased anti-tumor activity in a murine model, compared to the non-permuted parent fusion protein (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* 55:3357-3363, 1995; Puri, R. K. et al., *Cellular Immunol.* 171: 80-86, 1996). Increased potency of the permuted molecule is believed to result from a reduction in steric interference between the IL-4 domain in the parent molecule and its receptor.

Steric hindrance is likely to be a concern for other chimeric proteins which interact with receptors through a relatively large area of their surface. The same issue also arises with bioconjugates, containing relatively small chemicals conjugated to proteins or other molecules in complex polymers (Rose, K. et al., *Molecular Immunology* 32: 1031-1037, 1995).

## Phage display methods

Display methods allow affinity selection of protein variants from a library of displayed proteins or peptides (Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994; Winter, G., *Drug Development Res.* 33: 71-89, 1994). Many biological entities can be used in display methodologies (so-called "genetic packages" for presentation, including bacterial and eukaryotic cells, various eukaryotic and prokaryotic viruses, and spores), but the most commonly used vehicles used for display are filamentous bacteriophage, as used herein. We envision the possibility that a genetic package other than the particular phage used here could be used to present libraries of permuteins, and if so, constitute essentially the same invention.

Foreign proteins are presented on the surface of a phage particle, and the gene encoding the foreign protein is encapsulated in the virion. Because they are linked by the phage particle, affinity isolation of the presented protein also leads to affinity isolation of the corresponding genes. Extremely large libraries of phage presented proteins are constructed and affinity screened very rapidly. From the standpoint of how quickly mutant proteins can be made and screened for activity, phage display is the most efficient mutagenesis technique currently available.

#### Functional properties of permuteins

Permuteins can have improved biological properties by acting through several mechanisms. The permutein acting on the same type of cell as its parent molecule, may have increased binding, or other action, by virtue of increased avidity. Dimers or higher order multimers of these proteins with themselves or other chemical groups, including proteins, can have increased efficacy or potency, or both.

15

5

10

Permuteins can also have improved therapeutic properties through a variety of mechanisms such as: (1) alterations in the overall on- or off-rates or K, or K, of the ligand(s) on the target cell; (2) activation or blockade of complementary receptor signaling pathways; and/or (3) more specific targeting of to the cell of interest. The permuteins may also possess a unique pharmacokinetic distribution and clearance profile (Dehmer et al., Circulation, 91, 2188-2194, 1995; Tanaka et al., Nature Medicine, 3, 437-442, 1997).

20

Permuteins can also have improved properties in vivo, compared to the two components individually, as a result of alterations in biodistribution or half-life. The improved properties can also result from the binding of the permutein to one or more of the receptors, pharmacokinetics, or uptake of the permutein is altered in a favorable manner.

30

25

Molecular biology approaches have traditionally been used to permute proteins (Horlick, R.A. et al., *Protein Engineering* 5: 427-431, 1992) although chemical approaches have been used to make small permuted proteins (Goldenberg, D. P. and T. E. Creighton, *J. Mol. Biol.* 165: 407-413, 1983). These approaches are relatively labor intensive, limiting the number of permuteins that can be generated and efficiently screened for the desired biological activities. Rapid methods of generating permuteins, coupled with efficient methods for screening are needed that will result in the identification of novel active molecules.

## Summary of the invention

The present invention is an improved method for generating permuteins (scanning permutagenesis) based on the display of proteins on bacteriophage surface proteins. Phage display is a powerful, yet convenient tool, traditionally used for mutagenesis and screening (Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994). Improvements to this technology allow the rapid generation and screening of libraries of permuteins. Variables, such as position of the new termini and the length and composition of peptide linkers can easily be varied to generate libraries of the desired diversity.

10

15

5

The present invention relates to methods of producing biologically-active circularly permuted proteins of the formula C¹-L¹-N¹, derived from a parent protein of the formula N¹-C¹, wherein C¹ is comprised of a segment derived from the carboxy portion of said parent protein; N¹ is comprised of a segment derived from the amino terminal portion of said parent protein; and L¹ is a chemical bond or a linker, linking C¹ to the amino terminus of L¹ and carboxy terminus of L¹ to the amino terminus of N¹; comprising the steps of: (a) making a series of circularly-permuted genes; (b) inserting said circularly-permuted genes into a display vector; (c) expressing said circularly-permuted genes such that the proteins encoded by said genes are presented on the surface of the display vector; (d) generating a library of display vectors presenting the expressed circularly permuted protein; (e) affinity-select the presenting display vectors with a target protein that can bind a biologically-active circularly-permuted protein; (f) isolate and analyze clones of selected display vectors to identify the presented circularly-permuted protein.

20

Preferably the method of making a series of circularly-permuted genes is selected from the group consisting of making a tandemly-repeated intermediate, total synthesis of a synthetic gene, assembly of a gene from synthetic oligonucleotides, DNA amplification, and limited digestion of a circular intermediate.

30

25

Preferably, the display vector is selected from the group consisting of bacteriophage display vectors, bacteria, and baculovirus vectors. Even more preferably the presentation vector is a bacteriophage. Even more preferably, the presentation vector is bacteriophage M13. Most preferably, the presentation vector is a bacteriophage M13 gene III vector.

10

15

20

25

30

Preferably the method of making a series of circularly permuted genes is a method of making a tandem repeat intermediate. Even more preferably circularly permuted genes are amplified from the repeat by gene amplification.

Preferably the method of affinity selection comprises the steps consisting of

(a) binding said presentation display vectors to a target protein; (b) eluting said display vectors; (c) amplifying said display vectors; and (d) biopanning a pool of said amplified display vectors.

Preferably, the length of  $C^1$  in the permutein is longer than the length of  $C^1$  in said parent protein. More preferably, the length of  $C^1$  in the permutein is shorter than the length of  $C^1$  in said parent protein. Most preferably, the length of  $C^1$  in the permutein is the same length as the length of  $C^1$  in said parent protein.

Preferably, the length of  $N^1$  in the permutein is longer than the length of  $N^1$  in said parent protein. More preferably, the length of  $N^1$  in the permutein is shorter than the length of  $N^1$  in said parent protein. Most preferably, the length of  $N^1$  in the permutein is the same length as the length of  $N^1$  in said parent protein.

The invention also contemplates circularly permuted proteins of the formula C¹-L¹-N¹ made by the method of scanning permutagenesis. Preferably, the DNA sequence encoding said linker L¹ is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 368.

Preferably, the circularly-permuted protein is the G-CSF receptor agonist domain of a species of mylepoietin (MPO). MPO is one member of a family of novel dual cytokine receptor agonists (McKearn, J.P., Myelorestorative activities of synthokine and myelopoietin. In Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996) which are amenable to manipulation by phage display (Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997; Lee, S.C., Phage presentation for cytokine engineering. IBC's Second International Conference on Display Technologies, 1997).

## Brief description of the figures

#### Figure 1. Schematic depiction of scanning permutagenesis

Plate A of Figure 1 shows the strategy to generate a scanning permutagenesis phage display library. A plasmid containing directly-repeated

tandem copies of the hG-CSF gene, for example, is constructed by standard methods. The tandem repeat plasmid is used as the template for PCR amplification of genes encoding permuted proteins. Each copy of the G-CSF gene is indicated in light gray (turquoise), and a DNA segment encoding a peptide linker is indicated in dark gray (red).

In individual PCR reactions, oligonucleotide primers that initiate PCR polymerization at the first nucleotide of a chosen codon of G-CSF, and directing polymerization to the end of the tandem construct specifying the carboxy end of the protein encoded on the template is annealed to the tandem template. Also, a second specific primer is also annealed to the template that initiates polymerization at the last nucleotide of the codon encoding the amino acid immediately preceding the codon where polymerization begins with first primer, and which directs polymerization in the opposite direction from that first primer. Amplification between these two primers produces a DNA segment encoding a permuted protein. For example, amplification between the primer indicated by a black arrow initiating at codon 2 and the primer indicated by the blue arrow and initiating at the codon before 2 (codon 1) produces an amplified gene encoding a permuted protein whose amino terminal residue is amino acid 2 of the native protein, and whose final amino acid is amino acid 1 of the native protein.

20

25

5

10

15

A linker peptide is present between the first and final amino acids of the parent protein (residues 1 and 174 in this example). A total of 174 individual amplifications would produce a complete collection of all permuted proteins of this example. More limited collections containing only a selected set of permuteins can be made, as well as more extensive collections made from multiple tandem template plasmids, each containing a different linker sequence between the first and last residues of the two directly repeated tandem gene sequences. The collection of amplified segments can then be inserted into a phagemid presentation vector by standard methods. Phagemid particles produced from these presentation constructs are the scanning permutagenesis phage display library.

30

35

Plate B of Figure 1 shows the affinity screening of a phage display library (See Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994; Winter, G., *Drug Development Res.* 33: 71-89, 1994). In this example, a hG-CSF scanning permutagenesis library as described in Figure 1A is screened using the hG-CSF receptor expressed on mammalian cells as the affinity reagent. In Figure 1B, individual presented proteins are indicated by the shaded circles or diamonds and the affinity reagent is indicated by the light gray (pink) rectangles. Presentation

10

15

20

25

30

library particles are exposed to affinity reagent, unbound particles are washed away, and receptor-bound particles are eluted. The eluted particles are amplified in *E. coli*, and the screening cycle is repeated. During any round of the screening cycle, the genes encoded (in the present example encoding permuted proteins) by the selected particles can be expressed and evaluated.

## Figure 2. Permuteins presented in the scanning permutagenesis library

Human G-CSF (ser17) protein is depicted as a string of circles, each circle corresponding to a single amino acid residue. Amino and carboxy ends of the protein are indicated. The amino acids of helical regions are indicated by medium gray balls, while the amino acids of inter-helical loops are indicated in light gray balls (See Hill et al., *Proc. Natl. Acad. Sci. USA* 90: 5167-5171, 1993). Amino ends of the permuteins made for presentation in the library are indicated in dark gray. Asterisks indicate the breakpoints of the presented permuteins which were isolated by affinity screening with cells expressing hG-CSF receptor as illustrated in 1B.

# Figure 3. Bioactivity of permuteins identified by affinity screening of the scanning permutagenesis library

Individual permuteins were expressed transiently in mammalian cells. Permeation molecules in the culture supernatants were quantitated by ELISA, and the proliferative activity of clones was determined using BAF-3-cells dependent on G-CSF for growth. The horizontal axis indicate concentration of protein and the vertical axis indicate incorporation of tritiated thymidine.

## Definitions ,

The following is a list of abbreviations and the corresponding meanings as used interchangeably herein:

g = gram(s)

mg = milligram(s)

ml and mL = milliliter(s)

RT = room temperature

ug and µg = microgram(s)

uL and µl = microliter(s)

The following is a list of definitions of various terms used herein:

The term "permutein" means a circularly-permuted protein: a protein in which the amino and carboxy ends of the parent protein are joined together by a peptide linker sequence of zero or more amino acids. The amino and carboxy ends of the permuted protein occur at amino acids within the parental sequence.

5

The terms "chemical ligation" and "conjugation" mean a chemical reaction which covalently links two similar or dissimilar functional groups together \_ intramolecularly or intermolecularly.

10

The term "peptide linker' means a compound which forms a carboxamide bond between two groups having one or more peptide linkages (CONH-) and serves as a connector for the propose of amelioration of the distance or space orientation between two molecules.

The term "native sequence" refers to an amino acid or nucleic acid sequence which is identical to a wild-type or native form of a gene or protein.

15

20

The terms "mutant amino acid sequence," "mutant protein", "variant protein", "mutein", or "mutant polypeptide" refer to a polypeptide having an amino acid sequence which varies from a native sequence due to amino acid additions, deletions, substitutions, or all three, or is encoded by a nucleotide sequence from an intentionally-made variant derived from a native sequence.

# Detailed description of the invention

# Determination of the amino and carboxyl termini of permuteins

The present invention encompasses circularly permuted-proteins of the formula C'-L'-N' prepared by phage display techniques. The polypeptide can be joined either directly or through a linker segment. The term "directly" defines permuteins in which the polypeptide ends are joined without a linker. Thus L' represents a chemical bond or a linker, preferably a polypeptide segment to which both C' and N' are joined, wherein C' is comprised of a segment derived from the carboxy portion of the parent protein and N' is comprised of a segment derived from the amino terminal portion of a parent protein represented by the general formula N'-C'. Preferably, N' and C' in the permuted protein C'-L'-N' are the same length as in the parent protein N'-C', but each may be independently shorter or longer depending on the desired structural characteristics of the permutein. Most commonly L' is a linear peptide in which C' and N' are joined by amide

30

25

10

15

20

25

30

bonds, linking  $C^1$  to the amino terminus of  $L^1$  and carboxy terminus of  $L^1$  to the amino terminus of  $N^1$ .

Additional peptide sequences may also be added to facilitate purification or identification of permuteins (e.g., poly-His). A highly antigenic peptide may also be added that would enable rapid assay and facile purification of the permuteins by a specific monoclonal antibody.

#### Determination of the linker

The linking group (L¹) is generally a polypeptide of between 1 and 500 amino acids in length. The linkers joining the two molecules are preferably designed to (1) allow the two molecules to fold and act independently of each other, (2) not have a propensity for developing an ordered secondary structure which could interfere with the functional domains of the two proteins, (3) have minimal hydrophobic characteristics which could interact with the functional protein domains and (4) provide steric separation of C¹ and N¹ such that C¹ and N¹ could interact simultaneously with their corresponding receptors on a single cell. Typically surface amino acids in flexible protein regions include Gly, Asn and Ser. Virtually any permutation of amino acid sequences containing Gly, Asn and Ser would be expected to satisfy the above criteria for a linker sequence. Other neutral amino acids, such as Thr and Ala, may also be used in the linker sequence. Additional amino acids may also be included in the linkers due to the addition of unique restriction sites in the linker sequence to facilitate construction of the multi-functional proteins.

Preferred  $L^1$  linkers of the present invention include sequences selected from the group of formulas:

(SEQ ID NO: 1) through SEQ ID NO: 268)

Other linkers are also contemplated by the invention. The present invention is, however, not limited by the form, size or number of linker sequences employed. The only requirement of the linker is that it does not functionally interfere with the folding and function of the individual molecules of the multifunctional protein.

#### Utility of permuteins

Permuteins of the present invention may exhibit useful properties such as having similar or greater biological activity when compared to a single factor or by

10

15

20

25

30

having improved half-life or decreased adverse side effects, or a combination of these properties.

Permuteins which have little or no activity maybe useful as antigens for the production of antibodies for use in immunology or immunotherapy, as probes or as intermediates used to construct other useful permuteins.

The permuteins of the present invention may have an improved therapeutic profile as compared to their parent molecules. For example, some permuteins of the present invention may have a similar or more potent activities relative to other compounds or proteins without having a similar or corresponding increase in side-effects. This is particularly true of multifunctional or fusion protein therapeutics, where permutation may relieve steric and other hindrances that impair the activity of the parent fusion molecules (see Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* 55:3357-3363, 1995, for examples).

A general utility of permuteins is in the area of nanoscale devices described alternatively as "nanobiological" or "nanobiotechnological." These are nanoscale devices containing both precise structure nanomaterials and biological functional components (such as proteins). Nanodevices have been the subject of several reviews (Lee, S.C., Trends in Biotechnology, 16: 239-240, 1998).

Nanobiological/nanobiotechnological devices generally contain proteins covalently coupled to polymers or other non-biological precise structure materials. Issues of steric and other interferences with protein activity are applicable to proteins in nanobiological/nanobiotechnological devices and are highly analogous to the issues with multifunctional/fusion proteins discussed above. Protein permutation is fully expected to offer a viable approach to deal with these considerations, just as it does in the case of fusion proteins (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al.,, *Cancer Res.* 55:3357-3363, 1995).

#### Examples

The following examples will illustrate the invention in greater detail, although it will be understood that the invention is not limited to these specific examples. Various other examples will be apparent to the person skilled in the art after reading the present disclosure without departing from the spirit and scope of

the invention. It is intended that all such other examples be included within the scope of the appended claims.

#### General Materials and Methods

General methods of cloning, expressing, and characterizing proteins are found in T. Maniatis et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, 1982, and references cited therein, incorporated herein by reference; and in J. Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd edition, Cold Spring Harbor Laboratory, 1989, and references cited therein, incorporated herein by reference.

10

5

Unless noted otherwise, all specialty chemicals were obtained from Sigma Chemical Company (St. Louis, MO). Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs (Beverly, MA) or Boehringer Mannheim (Indianapolis, IN).

#### Strains, plasmids, and bacteriophage

15

20

25

The bacterial strains used in these studies are listed in Table 1. Plasmids and bacteriophage used or constructed in this study are listed in Tables 2 and 3, respectively.

Phage and phagemid stocks were made and manipulated as described (Kay, B.K., Winter, J., and McCafferty, J., Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996; Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997).

#### Transformation of E. coli strains

MD) and TGI (Amersham Corp., Arlington Heights, IL) are used for transformation of ligation reactions and are the hosts used to prepare plasmid DNA for transfecting mammalian cells. E. coli strains, such as JM101 (Yanisch-Perron et al., Gene, 33: 103-119, 1985) and MON105 (Obukowicz et al., Appl. and Envir. Micr., 58: 1511-1523, 1992) can be used for expressing the multi-functional

E. coli strains (Table 1), such as DH5α™ (Life Technologies, Gaithersburg,

proteins of the present invention in the cytoplasm or periplasmic space.

DH5α<sup>TM</sup> Subcloning efficiency cells are purchased as competent cells and are ready for transformation using the manufacturer's protocol, while both E. coli

strains TG1 and MON105 are rendered competent to take up DNA using a CaCl,

30

3NSDOCID: <WO

10

15

20

25

30

35

method. Typically, 20 to 50 mL of cells are grown in LB medium (1% Bactotryptone, 0.5% Bacto-yeast extract, 150 mM NaCl) to a density of approximately 1.0 optical density unit at 600 nanometers (OD600) as measured by a Baush & Lomb Spectronic spectrophotometer (Rochester, NY). The cells are collected by centrifugation and resuspended in one-fifth culture volume of CaCl<sub>2</sub> solution (50 mM CaCl<sub>2</sub>, 10 mM Tris-Cl, pH7.4) and are held at 4°C for 30 minutes. The cells are again collected by centrifugation and resuspended in one-tenth culture volume of CaCl<sub>2</sub> solution. Ligated DNA is added to 0.2 mL of these cells, and the samples are held at 4°C for 30-60 minutes. The samples are shifted to 42°C for two minutes and 1.0 mL of LB is added prior to shaking the samples at 37°C for one hour. Cells from these samples are spread on plates (LB medium plus 1.5% Bacto-agar) containing either ampicillin (100 micrograms/mL, ug/mL) when selecting for ampicillin-resistant transformants, or spectinomycin (75 ug/mL) when selecting for spectinomycin-resistant transformants. The plates are incubated overnight at 37°C.

Colonies are picked and inoculated into LB plus appropriate antibiotic (100 ug/mL ampicillin or 75 ug/mL spectinomycin) and are grown at 37°C while shaking.

#### DNA isolation and characterization

DNA constructs were made and propagated in *E. coli* using standard molecular biology techniques (Sambrook, J. et al., *Molecular Cloning, A Laboratory Manual*, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory, 1989).

Plasmid DNA can be isolated by a number of different methods and using commercially available kits known to those skilled in the art. Plasmid DNA is isolated using the Promega Wizard<sup>TM</sup> Miniprep kit (Madison, WI), the Qiagen QIAwell Plasmid isolation kits (Chatsworth, CA) or Qiagen Plasmid Midi or Mini kit. These kits follow the same general procedure for plasmid DNA isolation. Briefly, cells are pelleted by centrifugation (5000 x g), the plasmid DNA released with sequential NaOH/acid treatment, and cellular debris is removed by centrifugation (10000 x g). The supernatant (containing the plasmid DNA) is loaded onto a column containing a DNA-binding resin, the column is washed, and plasmid DNA eluted. After screening for the colonies with the plasmid of interest, the E. coli cells are inoculated into 50-100 ml of LB plus appropriate antibiotic for overnight growth at 37°C in an air incubator while shaking. The purified plasmid DNA is used for DNA sequencing, further restriction enzyme digestion, additional

10

15

20

25

30

subcloning of DNA fragments and transfection into E. coli, mammalian cells, or other cell types.

#### Sequence confirmation

DNA sequence analysis was performed using the Genesis 2000 DNA analysis system using standard methods (Prober et al., Science 238: 336-341, 1987).

Purified plasmid DNA is resuspended in dH<sub>2</sub>O and its concentration is determined by measuring the absorbance at 260/280 nm in a Bausch and Lomb Spectronic 601 UV spectrometer. DNA samples are sequenced using ABI PRISMTM DyeDeoxy™ terminator sequencing chemistry (Applied Biosystems Division of Perkin Elmer Corporation, Lincoln City, CA) kits (Part Number 401388 or 402078) according to the manufacturer's suggested protocol usually modified by the addition of 5% DMSO to the sequencing mixture. Sequencing reactions are performed in a DNA thermal cycler (Perkin Elmer Corporation, Norwalk, CT) following the recommended amplification conditions. Samples are purified to remove excess dye terminators with Centri-Sep™ spin columns (Princeton Separations, Adelphia, NJ) and lyophilized. Fluorescent dye labeled sequencing reactions are resuspended in deionized formamide, and sequenced on denaturing 4.75% polyacrylamide-8M urea gels using ABI Model 373A and Model 377 automated DNA sequencers. Overlapping DNA sequence fragments are analyzed and assembled into master DNA contigs using Sequencher DNA analysis software (Gene Codes Corporation, Ann Arbor, MI).

## Expression of permuted proteins in mammalian cells

To obtain sufficient protein for analysis of the bioactivity of individual MPO molecules containing cphG-CSFs, DNA segments containing individual affinity-selected MPO: cphGCSFs were subcloned into a mammalian expression vector, and expressed transiently in BHK cells as described below.

The BHK-21 cell line can be obtained from the ATCC (Rockville, MD). The cells are cultured in Dulbecco's modified Eagle media (DMEM/high-glucose), supplemented to 2 mM (mM) L-glutamine and 10% fetal bovine serum (FBS). This formulation is designated BHK growth media. Selective media is BHK growth media supplemented with 453 units/mL hygromycin B (CalBiochem, San Diego, CA). The BHK-21 cell line was previously stably transfected with the HSV transactivating protein VP16, which transactivates the IE110 promoter found on

10

15

20

the plasmid pMON3359 and pMON3633 and the IE175 promoter found in the plasmid pMON3360B (Hippenmeyer, P.J. and Pegg, L.E., Curr. Opin. Biotechnol. 6: 548-552, 1995). The VP16 protein drives expression of genes inserted behind the IE110 or IE175 promoter. BHK-21 cells expressing the transactivating protein VP16 are designated BHK-VP16. The plasmid pMON1118 expresses the hygromycin resistance gene from the SV40 promoter (Highkin et al., Poultry Sci., 70: 970-981, 1991). A similar plasmid, pSV2-hph, is available from ATCC.

BHK-VP16 cells are seeded into a 60 millimeter (mm) tissue culture dish at 3 x 10<sup>5</sup> cells per dish 24 hours prior to transfection. Cells are transfected for 16 hours in 3 mL of "OPTIMEM"<sup>TM</sup> (Gibco-BRL, Gaithersburg, MD) containing 10 ug of plasmid DNA containing the gene of interest, 3 ug hygromycin resistance plasmid, pMON1118, and 80 ug of Gibco-BRL "LIPOFECTAMINE" per dish. The media is subsequently aspirated and replaced with 3 mL of growth media. At 48 hours post-transfection, media from each dish is collected and assayed for activity (transient conditioned media). The cells are removed from the dish by trypsin-EDTA, diluted 1:10, and transferred to 100 mm tissue culture dishes containing 10 mL of selective media. After approximately 7 days in selective media, resistant cells grow into colonies several millimeters in diameter. The colonies are removed from the dish with filter paper (cut to approximately the same size as the colonies and soaked in trypsin/EDTA) and transferred to individual wells of a 24 well plate containing 1 mL of selective media. After the clones are grown to confluence, the conditioned media is re-assayed, and positive clones are expanded into growth media.

## Affinity selection and screening of phagemids

25

30

Affinity reagent used for the identification of functional MPO molecules containing cphG-CSF (MPO: cphG-CSF) species from the library were BHK cells expressing the hG-CSF receptor on their surface. The library pool was subjected to iterative affinity selection (four rounds) against BHK cells expressing the h-GCSF receptor using previously described techniques (Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997). Between rounds of selection, phage eluted from the affinity reagent were amplified in E. coli (Kay, B.K. J. Winter, and J. McCofferty, Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996).

10

15

20

**25** .

30

#### Expression of proteins in E. coli

When large-scale quantities of recombinant protein are desirable for structure-function studies, DNA segments containing individual affinity-selected MPO:cphGCSFs are subcloned into any of a variety of bacterial plasmid expression vectors, and expressed as a cytoplasmic product or as a secreted protein in *E. coli*.

E. coli strain MON105 or JM101 harboring the plasmid of interest are grown at 37°C in M9 plus casamino acids medium with shaking in an air incubator Model G25 from New Brunswick Scientific (Edison, NJ). Growth is monitored at  $\mathrm{OD}_{\infty}$  until it reaches a value of 1.0 at which time nalidixic acid (10 mg/mL) in 0.1 N NaOH is added to a final concentration of 50 µg/mL, for cultures containing plasmids with the E. coli recA promoter driving expression of the recombinant gene. IPTG is used in place of nalidixic acid, as a chemical inducer to facilitate expression from plasmids containing the lac promoter or hybrid lac promoters. The cultures are then shaken at 37°C for three to four additional hours. A high degree of aeration is maintained throughout the culture period in order to achieve maximal production of the desired gene product. The cells are examined under a light microscope for the presence of inclusion bodies (IB). One mL aliquots of the culture are removed for analysis of protein content by boiling the pelleted cells, treating them with reducing buffer and electrophoresis via SDS-PAGE (see Maniatis et al., "Molecular Cloning: A Laboratory Manual", 1982). The culture is centrifuged (5000 x g) to pellet the cells.

#### **Isolation of Inclusion Bodies**

The cell pellet from a 330 mL E. coli culture is resuspended in 15 mL of sonication buffer (10 mM 2-amino-2-(hydroxymethyl) 1,3-propanediol hydrochloride (Tris-HCl), pH 8.0 + 1 mM ethylenediaminetetraacetic acid (EDTA). These resuspended cells are sonicated using the microtip probe of a Sonicator Cell Disruptor (Model W-375, Heat Systems-Ultrasonics, Inc., Farmingdale, New York). Three rounds of sonication in sonication buffer followed by centrifugation are employed to disrupt the cells and wash the inclusion bodies (IB). The first round of sonication is a 3 minute burst followed by a 1 minute burst, and the final two rounds of sonication are for 1 minute each.

#### Purification

The folded proteins can be affinity-purified using affinity reagents such as monoclonal antibodies or receptor subunits attached to a suitable matrix.

10

15

20

25

30

Purification can also be accomplished using any of a variety of chromatographic methods such as: ion exchange, gel filtration or hydrophobic chromatography or reversed phase HPLC. These and other protein purification methods are described in detail (Methods in Enzymology, Volume 182 "Guide to Protein Purification" edited by Murray Deutscher, Academic Press, San Diego, California, 1990).

#### **Protein Characterization**

The purified protein is analyzed by RP-HPLC, electrospray mass spectrometry, and SDS-PAGE. The protein quantitation is done by amino acid composition, RP-HPLC, and Bradford protein determination. In some cases tryptic peptide mapping is performed in conjunction with electrospray mass spectrometry to confirm the identity of the protein.

## Baf-3/G-CSF receptor assay

Briefly, the mouse lymphoid cell line Baf3 was transfected with human granulocyte colony stimulating factor receptor (hG-CSFR) cDNA. Stable clones of Baf3 which expressed the G-CSFR and proliferated in the presence of hG-CSF were isolated and used to investigate the activity of human G-CSF receptor agonists without the influence of other human cytokine receptor responses.

The cDNA encoding hG-CSFR (a gift from Dr. Daniel C. Link (Washington University, St. Louis, MO) was released from the plasmid pEMCV.Sralpha as a HindIII/EcoRI (5' to 3') fragment, gel-purified, and inserted into the mammalian cell expression plasmid pcDNA3 (Invitrogen, San Diego, CA). This plasmid contains enhancer-promoter sequences from the immediate early gene of the human cytomegalovirus (CMV), a bovine growth hormone polyadenylation signal and transcription termination sequences, a neomycin resistance gene is present for the selection of G418 stable cell clones, and an ampicillin resistance gene for selection in E. coli. Ligation mixtures were transformed into E. coli strain TG1 [delta (lac-pro), supE, thi, hsdD5/F'(traD36, proA'B', lacI', lacZdeltaM15] and plasmid DNA was purified using a Qiagen Midiprep Plasmid Kit. The structure of plasmid DNAs containing hG-CSFR were confirmed by restriction enzyme analysis and by automated DNA sequence analysis using an ABI sequencing machine. One of several plasmids with the correct structure was selected and given the designation pMON30298.

Baß cells, maintained in complete growth medium (RPMI 1640 supplemented to 10% FBS and 10% Wehi 3B supernatant as a source for mouse IL-

10

15

20

25

30

35

3), were seeded at a subconfluent cell density of 10^5 cells/ml in growth media (RPMI 1640 5% FBS; 2 mM L-glutamine) the day prior to the electroporation. The cells were collected and rinsed twice in 10 ml serum-free RPMI 1640. The cells were diluted to 10^6/ml in serum-free RPMI and 1 ml was placed into each electroporation chamber (Gibco/BRL #1608AJ). 50 ug of plasmid DNA was added to each chamber and the chambers were incubated on ice for 30 minutes prior to electroporation. The cells were electroporated on ice at a capacitance of 800 uF, 400V, fast charge, and low ohms in a BRL CellPorator. The cells were immediately removed from the chambers and placed into 10 cm dishes containing 10 ml of growth medium. The cells were allowed to recover for 48 hr in growth media prior to selection.

After the recovery period, the cells were pelleted at 1000 rpm for 10 minutes, and resuspended into 10 ml of selection medium (growth medium containing 800 ug/ml G418 sulfate (Gibco/BRL). The cells were kept in selection media, being passaged twice weekly, until only a few viable cells could be seen in the mock transfected control cell dishes (approximately 2 weeks). After an additional 2 weeks in selection media, the cells which had been electroporated with the hG-CSFR cDNA had grown to a cell density which allowed them to be tested for proliferation in the presence of hG-CSF (Fukunaga, R. et al., EMBO J. 10 (10): 2855-2865, 1991).

The cell proliferation assay conditions are as follows: Briefly, 25,000 cells were plated in a microtiter 96 well plate with or without cytokine in IMDM medium supplemented with BSA (50 ug/ml), human transferrin (100 ug/ml), lipid (50 ug/ml) 2-mercaptoethanol (50 uM final concentration). Each well was incubated with 0.5 uCi of <sup>3</sup>H-thymidine (16 hours) and the incorporated radioactivity was measured. Triplicate wells containing Baf3 cells were set up with 4 nM hG-CSF, 4 nM mIL-3 or media only control. Samples of different permuted proteins were tested in each assay.

# Example 1: Construction of a permutein library without a linker region

Figure 1 shows a schematic of scanning permutagenesis. A plasmid construct comprising a tandem repeat of the modified human granulocyte colony stimulating factor (hG-CSF with a serine for amino acid 17) gene joined by a sequence (GCCGG, termed a zero order linker) was generated and subcloned into the plasmid pACYC177 (Chang, A.C.Y. and S.N. Cohen, *J Bacteriol*. 134: 1141-1156, 1978) using standard molecular biology methods (Sambrook, J. et al.,

Molecular Cloning, A Laboratory Manual, 2<sup>nd</sup> edition, Cold Spring Harbor Press, New York, 1989). The resultant plasmid construct (pMON15978) was linearized by restriction digestion (Smal) and used as a template for PCR amplification of circularly permuted hG-CSF (cphG-CSF) genes, following the method of Horlich (Horlick, R.A. et al., Protein Engineering 5: 427-431, 1992. For purposes of this demonstration of the scanning permutagenesis technique, we chose to make a limited permutein library rather than one containing every possible cphG-CSF. Figure 2 shows the position of the new amino termini for each new cphG-CSF.

10

15

5

vector pCANTAB 5E (Pharmacia Biotech,) such that they were expressed as a part of a MPO species (Feng, Y., N. R. Staten, C. M. Baum, N. L. Summers, M. Caparon, S. C. Bauer, L. Zurfluh, J. P. McKearn, B. K. Klein, S. C. Lee, C. A. McWherter. 1997. Multi-functional hematopoeitic receptor agonists. World Patent Application WO 97/12985) which was in turn fused to the amino end of the phage geneIII product. The presented fusion protein contained, starting from its amino terminus, a hIL-3 receptor agonist, cphG-CSF, and the phage gene III product. The juncture between the presented protein and the gene III product was as previously described (Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29,

Individual cphG-CSF genes were inserted into phagemid presentation

20

25

1997).

After confirmation of the structure of each phagemid construct, phagemid particles were produced for each individual cphG-CSF-presenting species (Merlin et al., 1997). Some of these lots of particles were used to individually define the affinity properties of specific presented cphG-CSF species in analytical biopanning experiments (Caparon, M. H. et al., *Molecular Diversity* 1: 241-246, 1996; Merlin, S. et al., *Applied Biochemistry and Biotechnology* 67: 15-29, 1997), but all of the phage particle lots were titered and equivalent numbers of transducing units of each particle preparation were pooled together to form the scanning permutagenesis library for hG-CSF in an MPO background. Figure 2 shows the MPO: cp hG-CSF species present in the library.

30

35

MPO: cphG-CSF 38/37 is an example of the nomenclature used to specify the identity of individual permuted proteins. It describes a MPO molecule containing a circularly permuted human G-CSF module (with the serine 17 substitution). The first amino acid of the cphG-CSF domain is amino acid 38 of the parent protein, and the last amino acid is residue 37 of the parent.

Example 2: Presentation and Affinity screening of the MPO: cphGCSF

library

MPO: cphG-CSF 38/37, is a full hG-CSF receptor agonist (McKearn, J.P., Myelorestorative activities of synthokine and myelopoietin. In Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996). It was presented on filamentous phage as a positive control to demonstrate that permuted proteins can be presented on the surface of phage particles and – affinity selected. After phagemid particles were produced from this construct, they were subjected to analytical biopanning using cells expressing the hG-CSF receptor as affinity reagent.

10

5

Table 1 shows that phage presented MPO: cphG-CSF 38/37 was affinity selected by cells expressing the hG-CSF. MPO: cphGCSF 38/37-GPIII fusion was expressed, secreted and assembled into phagemid particles, and could be affinity selected by the hG-CSF receptor. Permutagenesis of a protein does not appear to impair its successful presentation.

15

Relative to typical phage display libraries, the complexities of cp libraries are low, containing perhaps hundreds to thousands of individuals. The demonstration library here contained about 50 distinct clones, as opposed to more typical phage libraries containing more than 10<sup>5</sup> individuals (reviewed in Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994).

20

37 randomly chosen selectants from round 1, and fewer from subsequent rounds (17, 11 and 14 were picked from rounds 2, 3 and 4, respectively) were chosen for sequence analysis. The identity of the MPO: cphG-CSFs identified in each round is shown in Figure 2.

25

A total of 14 MPO: cphGCSF species were identified from the output of affinity selection (Figure 2). Most of the MPO: cphGCSF species identified from the library had new carboxy and amino termini in loop segments (9 of 14 permuteins identified), rather than in clearly defined secondary structures (See Hill et al., 1993 for the hG-CSF structure). Five selectants had termini within helical domains of hG-CSF (MPO: cphG-CSFs 13/12, 19/18, 71/70, 123/122 and 159/158). For three of these molecules (MPO: cphG-CSFs 13/12, 71/70 and 123/122) their new ends lie at the outermost ends of helices, and therefore perturbation of secondary structure caused by these permuteins may be minimal. However, MPO: cphG-CSF 19/18 and MPO: cphG-CSF 159/158 have new termini well within helix 1 and helix 4 of hG-CSF, respectively.

30

10

15

20

25

30

These data parallel the observations of Graf and Schachman, who developed a limited DNase I digestion method for "random" permutagenesis (Graf, R. and H. K. Schachman, *Proc Natl Acad Sci USA* 93:11591-11596, 1996). They identified two permutein species of aspartate transcarbamoylase that introduced new amino and carboxy ends into secondary structural domains and that retained biological activity. In their work, the majority of permuteins introducing ends into secondary structures (5/7 identified) were significantly diminished in activity. In contrast, we found a several permuteins that introduced helical breaks retained activity (See Below). The method used by Graf and Schechman frequently introduces point mutations, small insertions and deletions into the permuted proteins, potentially complicating the analysis of the effects of permutagenesis.

## Example 3: Biological activity of MPO: cphG-CSFs selected from the cp phage library

To obtain sufficient protein for analysis of the bioactivity of individual MPO molecules containing cphG-CSFs, DNA segments containing individual affinity-selected MPO: cphGCSFs were subcloned into a mammalian expression vector, and expressed transiently in BHK cells as described above.

The MPO: cphG-CSFs isolated from biopanning were all expressed transiently in mammalian cells and the amount of MPO: cphG-CSF in each supernatant was determined by sandwich hII-3 ELISA (Olins P.O. et al., *J. Biol. Chem.* 270: 23754-23760, 1995). The quantitated supernatants were then assayed for G-CSF receptor agonist activity in a Baf-3/G-CSF receptor assay (Figure 3, Table 4).

All but one of the transiently expressed MPO: cphG-CSF proteins exhibited G-CSF activity equivalent to or slightly better than that of the parent MPO molecule, including those MPO: cphG-CSFs with new carboxy and amino ends within helixes. The permutein encoded by pMON16021 with a breakpoint between positions 48 and 49 did not exhibit activity in the G-CSF-dependent proliferation assay. These data suggest that most of the proteins isolated from the library are competent to bind the hG-CSF receptor and produce a proliferation signal.

All references, patents, or applications cited herein are incorporated by reference in their entirety, as if written herein.

#### References

Buchwalder, A. et al., Biochemistry 31:1621-1630, 1992.

Caparon, M. H. et al., Molecular Diversity 1: 241-246, 1996.

Chang, A.C.Y. and S.N. Cohen, J Bacteriol. 134: 1141-1156, 1978.

5 Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994.

Feng et al., J. Mol. Biol. 259: 524-551, 1996.

Feng, Y., N. R. Staten, C. M. Baum, N. L. Summers, M. Caparon, S. C. Bauer, L. Zurfluh, J. P. McKearn, B. K. Klein, S. C. Lee, C. A. McWherter. 1997. Multifunctional hematopoeitic receptor agonists. World Patent Application WO 97/12985.

Fukunaga, R., Ishizaka-Ikeda, E., Pan, C. X., Seto, Y., and Nagata, S. Functional Domains of the Granulocyte Colony-Stimulating Factor Receptor. *EMBO J.* 10 (10):2855-2865, 1991.

Goldenberg, D. P. and T. E. Creighton, J. Mol. Biol. 165: 407-413, 1983.

15 Graf, R. and H. K. Schachman, Proc Natl Acad Sci USA 93:11591-11596, 1996.

Hahn, M. et al., Proc Natl Acad Sci USA 91: 10417-10421, 1994.

Highkin et al., Poultry Sci., 70: 970-981, 1991.

Hill et al., Proc. Natl. Acad. Sci. USA 90: 5167-5171, 1993.

Hippenmeyer, P.J. and L.E. Pegg, Curr. Opin. Biotechnol. 6: 548-552, 1995.

Horlick, R.A. et al., Protein Engineering 5: 427-431, 1992.

Jelinski, L.W. Biologically related aspects of nanoparticles, nanostructured materials and nanodevices. *In* "WTEC workshop on Global Assessment of R &D Status and Trends in Nanoparticles, Nanostructured Materials and Nanodevices", International Technology Research Institute, Loyola College, Baltimore, MD., S.

25 C., 1998.

001800541 1

Johnson, J. and F. M. Raushel, Biochemistry 35: 10223-10233, 1996.

Kay, B.K. J. Winter, and J. McCofferty, Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996.

Koebnik, R. and L. Kramer, J. Mol. Biol. 250: 617-626, 1995.

Kreitman, R. J. et al. Cytokine 7(4): 311-318, 1995.

5 Kreitman, R. J. et al., Proc Natl Acad Sci USA 91: 6889-6893, 1994.

Kreitman, R. J. et al., Cancer Res. 55:3357-3363, 1995.

Lee, S.C. Biotechnology for Nanotechnology. *Trends in Biotechnology*, **16**: 239-240, 1998.

Lee, S.C., Phage presentation for cytokine engineering. IBC's Second International Conference on Display Technologies, 1997.

Lin, X. et al., Protein Science 4: 159-166, 1995.

Lowman and Wells, J. Mol. Biol. 234: 564-578, 1993.

Luger et al., Protein Engineering 3: 249-258, 1990.

Luger et al., Science 243: 206-210, 1989.

Maniatis, T. et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982.

McKearn, J.P. Myelorestorative activities of synthokine and myelopoietin. In Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996.

20 Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997.

Mullins, L. S. et al., J Am. Chem Soc. 116: 5529-5533, 1994.

Murray Deutscher (ed), Methods in Enzymology, Volume 182 "Guide to Protein Purification," Academic Press, San Diego, California, 1990.

Obukowicz, M. et al., Appl. and Envir. Micr., 58: 1511-1523, 1992.

25 Olins P.O. et al., J. Biol. Chem. 270: 23754-23760, 1995.

Prober, J.M. et al., Science 238: 336-341, 1987.

Protosova, N. Y. et al., Protein Engineering 7: 1373-1377, 1994.

Puri, R. K. et al., Cellular Immunol. 171: 80-86, 1996.

Rose, K. et al., Molecular Immunology 32: 1031-1037, 1995.

Sambrook, J. et al., Molecular Cloning, A Laboratory Manual, 2<sup>nd</sup> edition, Cold Spring Harbor Press, New York, 1989.

Shortle, D. and J. Sondek, Curr. Opin. Biotechnol. 6: 299-305.

Smith, G. P., Curr. Opin. in Biotechnol. 2: 668-673, 1991.

Smith, G. P., Science 228: 1315-1317, 1985.

Sondek, J. and D. Shortle, Proteins 7: 387-393, 1990.

10 Thomas, J. W. et al., Proc Natl Acad Sci USA 92: 3779-3783, 1995.

Winter, G., Drug Development Res. 33: 71-89, 1994.

Yang, Y. R. and H. K. Schachman, Proc Natl Acad Sci USA 90: 11980-11984, 1993.

Yanisch-Perron et al., Gene, 33: 103-119, 1985.

Zhang, T. et al., Biochemistry 32: 12311-12318, 1993.

## Tables

Table 1: Circularly permuted proteins

Protein	Reference
Enzymes	
T4 lysozyme	Zhang et al., Biochemistry 32:12311-12318 (1993);
	Zhang et al., <i>Nature Struct. Biol.</i> 1:434-438 (1995)
dihydrofolate reductase	Buchwalder et al., <i>Biochemistry</i> <b>31</b> :1621-1630 (1994);
	Protasova et al., <i>Prot. Eng.</i> 7:1373-1377 (1995)
ribonuclease T1	Mullins et al., J. Am. Chem. Soc. 116:5529-5533 (1994);
·	Garrett et al., Protein Science 5:204-211 (1996)
Bacillus β-glucanase	Hahn et al., Proc. Natl. Acad. Sci. U.S.A. 91:10417- 10421 (1994)
aspartate transcarbamoylase	Yang and Schachman, Proc. Natl. Acad.
aspartate transcarbanioyanse	Transcarbamoylase Sci. U.S.A. 90:11980-11984 (1993)
phosphoribosyl anthranilate	Luger et al., Science 243:206-210 (1989);
isomerase	Luger et al., Prot. Eng. 3:249-258 (1990)
pepsin/pepsinogen	Lin et al., Protein Science 4:159-166 (1995)
glyceraldehyde-3-phosphate dehydrogenase	Vignais et al., Protein Science 4:994-1000 (1995)
ornithine decarboxylase	Li & Coffino, Mol. Cell. Biol. 13:2377-2383 (1993)
yeast phosphoglycerate	Ritco-Vonsovici et al., Biochemistry 34:16543-
dehydrogenase	16551 (1995)
Enzyme Inhibitor	
basic pancreatic trypsin inhibitor	Goldenberg & Creighton, J. Mol. Biol. 165:407-413 (1983)
Cytokines	
interleukin-1ß	Horlick et al., Protein Eng. 5:427-431 (1992)
interleukin-4	Kreitman et al., Cytokine 7:311-318 (1995)

Tyrosine Kinase Recognition

Domain

α-spectrin SH3 domain

Viguera et al., J. Mol. Biol. 247:670-681 (1995)

Transmembrane Protein

omp A

Koebnik & Krämer, J. Mol. Biol. 250:617-626

(1995)

Chimeric Protein

interleukin-4-Pseudomonas exotoxin fusion molecule

Kreitman et al., Proc. Natl. Acad. Sci. U.S.A.

91:6889-6893 (1994)

Table 2: Strains

Designation	Description or Genotype	Reference/Source
DН5α™	F, phi80 dlacZdeltaM15, delta(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17 (rk,mk*), phoA, supE44, lambda-, thi-1, gyrA96, relA1	Life Technologies, Rockville, Maryland –
JM101 (ATCC# 33876)	delta (pro lac), supE, thi, F'(traD36, proA*B*, lacI*, lacZdeltaM15)	Yanisch-Perron et al., <i>Gene</i> , 33: 103-119, 1985
MON105 (ATCC# 55204)	F, lambda-,IN (rrnD, rrnE)1, rpoD <sup>*</sup> , rpoH358	Obukowicz et al., <i>Appl. and Envir. Micr.</i> , 58: 1511-1523, 1992
MON208	W3110 rpoH358, lacI <sup>9</sup> , ompT::kan	Alan Easton
TG1	delta(lac-pro), supE, thi-1, hsdD5/F'(traD36, proA*B*, lacIq, lacZdeltaM15)	Amersham Corp., Arlington Heights, Illinois
W3110	IN (rrnD-rrnE)1, rph1	Lab collection

Table 3: Plasmids

Plasmid	SEQ ID NO.	Selectable Marker	Description	Source
pACYC177		KanR AmpR	Plasmid with multiple cloning sites and two selectable markers	Chang, A.C.Y. and S.N. Cohen, J Bacteriol. 134: 1141- 1156, 1978
pMON15978		AmpR	Plasmid construct comprising a tandem repeat of the modified human granulocyte colony stimulating factor (hG-CSF with a serine for amino acid 17) gene joined by a sequence (GCCGG, termed a zero order linker), subcloned into the plasmid pACYC177	This work
pCANTAB 5E		AmpR	Phage display vector containing lac promoter operably linked to fd gene 3 signal sequence, a linker region, an E tag, and an fd gene 3 structural gene all cloned into the vector backbone of pUC119 containing ColE1 ori, the beta lactamase resistance gene, and an M13 ori.	Pharmacia Biotech, Piscataway NJ
pMON16016		$_{ m Amp}^{ m R}$	Phagemid presentation vector pCANTAB 5E derivation containing inserted individual cphG-CSF gene such that it was expressed as a part of an	This work

MPO species, fused in turn to the amino terminus end of the phage geneIII product. The first amino acid of the cphG-CSF domain is amino acid 1 of the parent, and the last amino acid is residue 174 of the parent. The zero order linker is attached at the carboxyl end of amino acid 174.

pMON16017

 $Amp^R$ 

Identical to pMON16016 except

This work

the first amino acid of the cphG-CSF domain is amino acid 3 of the parent, and the last amino acid is residue 2 of the parent.

pMON16029

 $\mathbf{Amp}^{\mathbf{R}}$ 

Identical to pMON16016 except This work

the first amino acid of the cphG-CSF domain is amino acid 7 of the parent, and the last amino acid is residue 6 of the parent.

pMON16030

 $Amp^R$ 

Identical to pMON16016 except

This work

the first amino acid of the cphG-CSF domain is amino acid 9 of the parent, and the last amino acid is residue 8 of the parent.

pMON16018

 $Amp^R$ 

Identical to pMON16016 except

This work

the first amino acid of the cphG-CSF domain is amino acid 11 of the parent, and the last amino acid is residue 10 of the parent.

pMON16019	$_{ m Amp}R$	Identical to pMONI6016 except the first amino acid of the cphG-CSF domain is amino acid 13 of the parent, and the last amino acid is residue 12 of the parent.	This work
pMON16031	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 15 of the parent, and the last amino acid is residue 14 of the parent.	This work
pMON16020	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 19 of the parent, and the last amino acid is residue 18 of the parent.	This work
pMON16032	Amp <sup>R</sup>	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 22 of the parent, and the last amino acid is residue 21 of the parent.	This work
pMON16033	$Amp^R$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 27 of the parent, and the last amino acid is residue 26 of the parent.	This work
pMON16034	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 31 of the parent, and the	This work

		last amino acid is residue 30 of the parent.	
pMON16035	$_{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 35 of the parent, and the last amino acid is residue 34 of the parent.	This work
pMON16036	${ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 37 of the parent, and the last amino acid is residue 36 of the parent.	This work
pMON16037	$^{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 38 of the parent, and the last amino acid is residue 37 of the parent.	This work
pMON16038	$_{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 39 of the parent, and the last amino acid is residue 38 of the parent.	This work
pMON16039	$_{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 43 of the parent, and the last amino acid is residue 42 of the parent.	This work
pMON16040	$_{\mathbf{Amp}^{\mathbf{R}}}$	Identical to pMON16016 except the first amino acid of the	This work

cphG-CSF domain is amino acid 45 of the parent, and the last amino acid is residue 44 of the parent.

last amino acid is residue 55 of

the parent.

Identical to pMON16016 except This work  $Amp^R$ **DMON16041** the first amino acid of the cphG-CSF domain is amino acid 47 of the parent, and the last amino acid is residue 46 of the parent. Identical to pMON16016 except  $Amp^R$ pMON16022 the first amino acid of the cphG-CSF domain is amino acid 49 of the parent, and the last amino acid is residue 48 of the parent. Identical to pMON16016 except This work  $Amp^R$ pMON16042 the first amino acid of the cphG-CSF domain is amino acid 51 of the parent, and the last amino acid is residue 50 of the parent. Identical to pMON16016 except This work  $Amp^R$ pMON16043 the first amino acid of the cphG-CSF domain is amino acid 53 of the parent, and the last amino acid is residue 52 of the parent. Identical to pMON16016 except This work  $\mathbf{Amp}^{\mathbf{R}}$ pMON16044 the first amino acid of the cphG-CSF domain is amino acid 56 of the parent, and the

ntiennin aan mioneai I

pMON16023	$_{ m Amp}{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 60 of the parent, and the last amino acid is residue 59 of the parent.	This work
pMON16045	$_{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 64 of the parent, and the last amino acid is residue 63 of the parent.	This work
pMON16024	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 67 of the parent, and the last amino acid is residue 66 of the parent.	This work
pMON16046	$_{\mathbf{Amp}}^{\mathbf{R}}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 69 of the parent, and the last amino acid is residue 68 of the parent.	This work
pMON16025	$_{ m Amp}$ R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 71 of the parent, and the last amino acid is residue 70 of the parent.	This work
pMON16047	$_{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 73 of the parent, and the	This work

pMON16048

pMON16049

pMON16050

 $\mathbf{Amp}^{\mathbf{R}}$ 

 $Amp^R$ 

 $Amp^R$ 

35		
	last amino acid is residue 72 of	
	the parent.	
	Identical to pMON16016 except	This work
	the first amino acid of the	
	cphG-CSF domain is amino	
	acid 84 of the parent, and the	-
	last amino acid is residue 83 of	
	the parent.	
	Identical to pMON16016 except	This work
	the first amino acid of the	
	cphG-CSF domain is amino	
	acid 98 of the parent, and the	
	last amino acid is residue 97 of	
	the parent.	•
	Identical to pMON16016 except	This work
	the first amino acid of the	
	cphG-CSF domain is amino	
	acid 100 of the parent, and the	
	last amino acid is residue 99 of	
	the parent.	
	Identical to pMON16016 except	This work
	the first amino acid of the	
	cphG-CSF domain is amino	
	acid 102 of the parent, and the	
	last amino acid is residue 101	
	of the parent.	
Ł	Identical to pMON16016 except	This work
	the first amino acid of the	
	cphG-CSF domain is amino	
	acid 112 of the parent, and the	

 $Amp^R$ pMON16051 pMON16052  $Amp^R$ acid 112 of the parent, and the last amino acid is residue 111 of the parent. Identical to pMON16016 except This work  $Amp^R$ pMON16053 the first amino acid of the

pMON16026

pMON16027

pMON16054

pMON16055

pMON16056

	30		
		cphG-CSF domain is amino acid 121 of the parent, and the last amino acid is residue 120 of the parent.	
AmpR		Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 123 of the parent, and the last amino acid is residue 122 of the parent.	This work
AmpR		Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 125 of the parent, and the last amino acid is residue 124 of the parent.	This work
AmpR		Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 133 of the parent, and the last amino acid is residue 132 of the parent.	This work
Amp <sup>R</sup>		Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 142 of the parent, and the last amino acid is residue 141 of the parent.	This work
AmpR		Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 143 of the parent, and the	This work

last amino acid is residue 142

of the parent.

-27

pMON16057	$\mathbf{Amp}^{\mathbf{R}}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 147 of the parent, and the last amino acid is residue 146 of the parent.	This work
pMON16028	Amp <sup>R</sup>	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 159 of the parent, and the last amino acid is residue 158 of the parent.	This work
pMON16058	Amp <sup>R</sup>	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 168 of the parent, and the last amino acid is residue 167 of the parent.	This work
pMON16059	$_{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 170 of the parent, and the last amino acid is residue 169 of the parent.	This work

RNSDOCID- ZWO DOTRADSA1 I

5

10

Table 4: Analytical biopanning

Before receptor*	After receptor*	Enrichment
1/6.6x10 <sup>4</sup>	1/6.5x10 <sup>-1</sup>	990-fold

<sup>\*</sup> Amp<sup>R</sup>/Kan<sup>R</sup> resistant colonies

Analytical biopanning shows that MPO molecules containing permuted hG-CSF domains can be presented and affinity selected in a hG-CSF receptor dependent fashion. A mixture of phagemids presenting MPO: cphG-CSF 38/37 (ampicillin resistant) and M13k07 (kanamycin resistant) were exposed to BHK cells with or without the hG-CSF receptor on their surface, washed and eluted from the cell surface. Eluted phage were introduced into *E. coli* and the transfected cells were plated on media containing kanamycin or ampicillin. The ratio of ampicillin resistant to kanamycin resistant particles were determined prior to and following exposure to receptor by counting resistant colonies.

Table 5: Activity of selected permuteins

Plasmid	Permutein breakpoint in G-CSF amino acid sequence	Activity in G-CSF- dependent proliferation assay
pMON16017	3/2	+
pMON16018	11/10	+
pMON16019	13/12	, +
pMON16020	19/18	+
pMON16021	49/48	<del>-</del>
pMON16022	60/59	+
pMON16023	67/66	+
pMON16024	69/68	+
pMON16025	71/70	* <b>+</b>
pMON16026	123/122	+
pMON16027	125/124	+
pMON16028	159/158	+

Table 6: SEQ ID Number/SEQ ID Name Correlation

520	SEQ ID	Sequence									
ID	Hame										
NO.					363	CCA	TTG	GGC	CCT	GCC	AGC
1.	FGS1	CCCCCCC	ACATG	TCT	ACA	TTG	GGC	CCT	GCC	AGC	TCC
2.	FGS2	CCCCCCC	ACATG	TCT .	CCA TTG	660	CCI	GCC	AGC	TCC	CTG
3.	FGS3	CCCCCCC	ACATG	TCT		CCI	GCC	AGC	TCC	CTG	CCC
٠4.	FGS4	CCCCCCC	ACATG	TCT	GGC	GCC.	AGC	TCC	CTG	CCC	CAG
5.	PGS5	CCCCCCC	ACATG	TCT	CCT		TCC	CIG	ccc	CAG	AGC
€.	FGS6	cececec	ACATG	TCT	GCC	AGC	CIC	.000	CAG	AGC	TTC
7.	FGS7	cececec	ACATG	TCT	AGC	100	CCC	CAG	AGC	TTC	CTG
8.	FGS8	CCCCCCC	ACATG	TCT	TCC	CTG	CAG	AGC	TTC	CTG	CTC
9.	FGS9	CCCCCCC	ACATG	TCT	CIG	CCC	AGC	110	CTG	CIC	AAG
10.	FGS10	CCCCCCC	ACATG	TCT	ccc	-CAG	TTC	CTG	CTC	AAG	TCT
11.	FGS11	CCCCCCC	ACATG	TCT	CAG	AGC TTC	CTG	CTC	AAG	TCT	TTA
12.	FGS12	CCCCCCC	ACATG	TCT	AGC	CIG	CIC	AAG	TCT	TTA	GAG .
13.	PGS13	CGCGCGC	ACATG	TCT	330	C.C.	AAG	TCT	TTA	GAG	CAA
14.	FGS14	CCCCCCC	ACATG	TCT	CLC	AAG	TCT	TTA	GAG	CAA	GTG
15.	FGS15	CGCCCCC	ACATG	TCT	CTC AAG	TCT	TTA	GAG	CAA	GTG	AGG
16.	FGS16	CCCCCCC	ACATG	TCT	TCT	TTA	GAG	CAA	GTG	AGG	AAG
17.	FGS17	CECECEC	ACATG	TCT		GAG	CAA	GTG	AGG	AAG	ATC .
18.	FGS18	CCCCCCC	ACATG	TCT	TTA GAG	CAA	GTG	AGG	AAG	ATC	CAG
19.	FG519	CCCCCCC	ACATG	TCT	CAA	GTG	AGG	AAG	ATC	CAG	GGC
20.	PG520	CGCGCGC	ACATG	TCT	GTG	AGG	AAG	ATC	CAG	GGC	GAT
21.	FGS21	CCCCCCC	ACATG	TCT		AAG	ATC	CAG	GGC	GAT	GGC
22.	FG522	cececec	ACATG	TCT	AGG	ATC	CAG	GGC	GAT	GGC	GCA
23.	FGS23	CCCCCCC	ACATG	TCT	AAG ATC	CAG	GGC	GÁT	GGC	GCA	GCG
24.	FGS24	CGCGCGC	ACATG	TCT	CAG	GGC	GAT	GGC	GCA	GCG	·CTC
25.	FGS25	CCCCCCC	ACATG		GGC	GAT	GGC	GCA	GCG	CTC	CAG
26.	FGS26	CCCCCCC	ACATG	TCT	GAT	GGC .	GCA	GCG	CIC	CAG	GAG
27.	FGS27	CCCCCCC	ACATG ACATG	TCT	GGC	GCA	GCG	CTC	CAG	GAG	AAG
28.	FGS28	CGCGCGC		TCT	GCA	GCG	CTC	CAG	GAG	AAG .	CTG
29.	PGS29	CCCCCCC	ACATG ACATG	TCT	GCG	CTC	CAG	GAG	AAG	CTG	TGT
30.	PGS30	CGCGCGC	ACATG	707	CTC	CAG	GAG	AAG	CTG	TGT	GCC
31.	PGS31	CGCGCGC	ACATG	TCT	CAG	GAG	AAG	CIG	TCT	GCC	ACC
32.	FGS32	CGCGCGC	ACATG	TCT	GAG	AAG	CTG	TGT	GCC	ACC	TAC
33.	PGS33	CGCGCGC	ACATG	TCT	AAG	CTS	TGT	GCC	ACC	TAC	AAG
34.	FGS34	CGCGCGC	ACATG	TCT	CIG	TCT	CCC	ACC	TAC	YYC	CTG
35.	PGS35	CCCCCCC	ACATG	TCT	TGT	GCC	ACC	TAC	AAG	CTG	1CC
36.	PGS36	CCCCCCC	ACATG	TCT	GCC	ACC	TAC	AAG	CIG	TGC	CAC
37.	PGS37	CCCCCCC	ACATG	TCT	ACC	TAC	AAG	CIG	TGC	CYC	-ccc
38.	PGS38	cececec	ACATG	TCT	TAC	AAG	CIG	TGC	CXC	ccc	GAG
39.	PGS39 PGS40	CGCGCGC	ACATG	TCT	AAG	CIG	TGC	CAC	CCC	· GAG	GAG
40.	PGS41	CGCGCGC	ACATG	TCT	CTG	TGC	CAC	CCC	GAG	GAG	CIG
41. 42.	FGS42	ccccccc	ACATG	TCT	TGC	CAC	ccc	GAG	GAG	CTG	GTG
43.	PGS43	CGCGCGC	ACATG	TCT	CAC	CCC.	GAG	GAG	CLC	GTG	CIG
44.	PGS44	cececec	ACATG	TCT	CCC	GAG	GAG	CTG	GTG	··CTG	CTC
45.	PGS45	CCCCCCC	ACATG	TCT	GAG	GAG	CIG	GTG	CTG	·CTC	GGA
46.	PGS46	CGCGCGC	ACATG	TCT	GAG	CIC	GTG .	CIG	CTC	GGA	CAC
47.	PGS47	CCCCCCC	ACATG	TCT	CIG	GTG	CTG	CTC	ÇGA	CAC	CTG
48.	FGS48	· ccccccc	ACATG	TCT	GTG	CIG	CIC	GGA	CAC	TCT	GGC
49.	PGS49	CGCGCGC	ACATG	TCT	CIG	CIC	GGA	CAC	TCT	CTG	ATC
50.	FGS50	CGCGCGC	ACATG	TCT	CIC	GGA	CAC	TCT	CTG	GGC ATC	-ccc
51.	FGS51	CCCCCCC	ACATG	TCI	GGA	CYC	TCT	CTG		CCC	100
52.	FGS52	CCCCCCC	ACATG	TCT	CAC	TCT	CTG	GGC	ATC	TGG	GCT
53.	FGS53	cececec	ACATG	<b>TCT</b>	TCT	CTG	GGC	ATC CCC	TGG	GCT	-ccc
54.	FGS54	CCCCCCC	ACATG	TCT	CTG	GGC	CCC	TGG	GCT	CCC	CIG
55 .	FGS55	CGCCCCCC	ACATG	TCT	GGC	ATC CCC	TGG	GCT	ccc	CTG	AGC
56.	<b>7</b> G256	cececec	ACATG	TCT	ATC		GCT	CCC	CIG	AGC	TCC
57.	PG557	CCCCCCC	ACATG	707	CCC	TGG	CCC	CTG	AGC	TCC	TGC
58.	FGS58	COCCCCC	ACATG	TCT	TCG	CCC	CTG	AGC	TCC	TGC	CCC
59.	FGS59	CCCCCCC	ACATG	TCT	CCI		AGC	TCC	TGC	ccc	AGC
€ů.	PGS60	CCCCCCC	ACREG	TCT	CCC	CTG AGC	TCC	TGC	222	AGC	CAG
61.	F3561	CGCGCGC	ACATG	10.1	CTG AGC	TCC	TGC	.000	AGC	CAG	GCC
62.		ccccccc	ACATG	TCT	TCC	TGC	CCC	AGC	CAG	GCC	CTG
63.		CGCGCGC	ACATG	TCT	TOC	ÇCC 10C	AGC	CAG.	GCC	CTG	CAG
64.		ccccccc	ACATG	TCT	CCC	AGC	CAG	GCC	CIG	CAG	CTG
65	PGS65	CCCCCCC	ACATG	TCT	AGC	CAG	GCC	CTG	CAG	CTG	GCA
66.		CGCGCGC	'ACATG	101	CAG	GCC	CTG	CAG	CTG	GCA	GGC
67.		CCCCCCC	ACATG	TCT TCT	GCC	CTG	CAG	CTG	GCA	GGC	TGC
68.		2200000	ACATG		CTG	CAG	CTG	GCA	GGC	760	TTG
69.		CCCCCCC	ACATG	TCT	CAG	276	GCA	GGC	TGC	170	AGC
~0.		CECECEC	ACATG	TCT	CVC	GCA	GCC	TCC	TTG	AGC	CAA
7:		2020202	ACATE	TCT	GCA	GGC	TGC	775	AGC	·CAA	CIC
72.		GGGGGG	ACATG		GGC	.100	776	AGC	CAA	CIC	-CAT
73 .		CCCCCCC	ACATG	TCT	760	775	AGC	CAA	CTC	CAT	ASC
74.		CCCCCCC	ACATG	701	125	A:32	CAA	CTC	CAT	AGC	GGC
75.		cececes	ACATG	707	NUC	TAA.	270	CAT	AGC	63:	CTT
76.		cececec	ACATG	TOT	CâÀ		·CAT	AGC	GGC	===	TTC
77.		0000000	ACATG ACATG	TCT	====	CAT	AGC	GSC	Cit	****	CIC
78.		CCCCCCC	ACATG	707	CAT	AGC	-G30	CTT	TTC	577	TAC
79.	. PGS79	ಂದಾರಾವ	WCVIO								

80.	FGS80	CCCCCCC	ACATG	TCT	AGC	CCC	CTT	777	CAC	TAC	CAG
61.	FGS81	CCCCCCC	ACATG	TCT	GCC	·CII	440	ctc	TAC	CAG	GGG
62.	FGSB2	CCCCCCC	ACATG	TCT	CII	3.10	CTC	TAC	CAG	GGG	CIC
83.	FGS83	CCCCCCC	ACATG	TCT	TTC	CTC	TAC	CAG	GCC	CTC	CIG
84.	PGS84	cececec	ACATG	TCT	CTC	TAC	CAG	CCC	-CIC	CTG	CAG
85.	FGS85	CCCCCCC	ACATG	TCT	TAC	CAG	GCC	CIC	·CTG	CAG	GCC
86.	FGS86	CCCCCCC	ACATG	TCT	CAG	GGG	CTC	CTG	·CAG	CCC	CTG
87.	FGS87	CCCCCCC	ACATG	TCT	CCC	CTC	CTG	-CAG	GCC	CIG	'GAA
	PGS88	CCCCCCC	ACATG	TCT	CTC	CTG	CAG	GCC	CIG	GAA	GGG
88		CCCCCCC	ACATG	TCT	CTG	CAG	GCC	CTG	GAA	GGG	ATA
89.	PGS89			TCT	CAG	GCC	CTG	GAA	GGG	ATA	TCC
90.,	FGS90	cccccc	ACATG	TCT	GCC	CTG	GAA	GGG	ATA	TCC	CCC
91.	FGS91	CCCCCCC	ACATG			GAA	GGG	ATA	TCC	CCC	GAG
92.	FGS92	CCCCCCC	ACATG	TCT	CTG			TCC	.000	GAG	TTG
93.	FGS93	CCCCCCC	ACATG	TCT	GAA	GGG	ATA				GGT
94.	FGS94	CCCCCCC	ACATG	TCT	GCC	ATA	TCC	CCC	GAG	TTG	
95.	PGS95	CCCCCCC	ACATG	TCT	ATA	1CC	ccc	GAG	TTG	GCT	CCC
96.	PGS96	CGCGCGC	ACATG	TCT	TCC	CCC	GAG	TIG	-GGT	CCC	ACC
97.	FGS97	CCCCCCC	ACATG	TCT	CCC	GAG	TTG	GGT	CCC	ycc	TIG
98.	FGS98	CGCGCGC	ACATG	TCT	GAG	TTG	GGT	CCC	ACC	TTG	GAC
99.	FGS99	cececec	ACATG	TCT	TTG	GGT	CCC	ACC	176	GAC	ACA
100.	FGS100	CGCGCGC	ACATG	TCT	CCT	CCC	ACC	TIG	GAC	ACA	CTG
101.	PGS101	CGCGCGC	ACATG	TCT	CCC	ACC	TTG	GAC	ACA	CTC	CAG
102.	FGS102	CGCGCGC	ACATG	TCT	ACC	TTG	GAC	ACA	-CL3	CAG	CIG
103.	FGS103	CGCGCGC	ACATG	TCT	TTG	GAC	ACA	CTG	CAG	CLC	GAC
104.	FGS104	CCCCCCC	ACATG	TCT	GAC	ACA	CTG	CAG	CIG	GAC	GTC
105.	FGS105	CCCCCCC	ACATG	TCT	ACA	CTG	CAG	CTG	GAC	GTC	GCC
106.	FGS106	CCCCCCC	ACATG	TCT	CTG	CAG	CTG	GAC	GTC	GCC	GAC
	PGS107	CCCCCCC	ACATG	TCT	CAG	CTC	GAC	GTC	GCC	GAC	TTT
107.	FGS107	CGCGCGC	ACATG	TCT	CTG	GAC	GTC	GCC	GAC	TIT	GCC
108.		CGCGCGC	ACATG	TCT	GAC	GTC	GCC	GAC	TIT	-GCC	ACC
109.	FGS109		ACATG	TCT	GTC	GCC	GAC	TTT	GCC	ACC	ACC
110.	FGS110	CGCGCGC		TCT	GCC	GAC	TTT	-622	' ACC	ACC	ATC
111.	FGS111	CGCGCGC	ACATG	TCT	GAC	TTT	GCC	ACC	ACC	ATC	TGG
112.	FGS112	cececec	ACATG		TTT	GCC	ACC	ACC	ATC	TGG	CAG
113.	FGS113	CGCGCGC	ACATG	TCT	GCC	ACC	ACC	ATC	TGG	CAG	CAG
114.	FGS114	cececec	ACATG	TCT				73G	CAG	CAG	ATG
115.	PGS115	CCCCCCC	ACATG	TCT	XCC	ACC	ATC			ATG	GAA
116.	FGS116	CCCCCCC	ACATG	TCT	YCC	ATC	TGG	CAG	CAG		
117.	FGS117	CCCCCCC	ACATG	TCT	ATC	TGG	CAG	CAG	ATG	GAA	GAA
118.	PGS118	CCCCCCC	ACATG	TCT	TGG	CAG	CAG	ATG	GAA	GAA	CTG
119.	PG5119	CCCCCCC	ACATG	TCT	CAG	CAG	ATG	GAA	GAA	CIG	GGA
120.	FGS120	CCCCCCC	ACATG	TCT	CAG	ATC	GAA	GAA	CTG	GGA	ATG
121.	PGS121	CGCGCGC.	<b>ACATG</b>	TCT	ATG	GAA	GAA	-CTG	GGA	YIG	GCC
122.	FGS122	-ccccccc.	ACATG	TCT	GAA	GAA	CIC	GGA	ATC	CCC	CCT
123.	PG5123	CCCCCCC	ACATG	TCT	GAA	CIG	GGA	ATG	CCC	CCT	GCC
124.	PGS124	CGCGCGC	ACATG	TCT	CIG	GGA	ATG	GCC -	CCT	GCC	-CTG
125.	FGS125	CGCGCGC	ACATG	TCT	GGA	ATG	GCC	-CCT	GCC	CIG	CAG
126.	FGS126	CCCCCCC	ACATG	TCT	ATG	GCC	CCT	GCC	CTG	CAG	222
127.	FGS127	cececec	ACATG	TCT	GCC	CCT	GCC	CTG	CAG	ccc	ACC
128.	FGS128	CCCCCCC	ACATG	TCT	CCT	GCC	CIG	CAG	CCC	ACC	CAG
129.	FGS129	cececec	ACATG	TCT	CCC	CIG	CAG	222	ACC	CAG	GGT
130.	PGS130	CGCGCGC	ACATG	TCT	CTG	CAG	CCC	ACC	CAG	GGT	GCC
131.	FGS131	cececec	ACATG	TCT	CAG	-ccc	ACC	CAG	GGT	GCC	ATG
132.	FGS132	CCCCCCC	ACATG	TCT	CCC	ACC	CAG	CST	· GCC	ATG	CCG
133.	FGS133	cececec	ACATG	TCT	ACC	CAG	CCT	GCC	ATG	·ccs	GCC
134.	FGS134	CGCGCGC	ACATG	101	CAG	GGT	GCC	ATG	ccs	-GCC	TTC
	FGS135	CGCGCGC	ACATG	TCT	GGT	GCC	ATG	ccs	GCC	TIC	GCC
125.			ACATG	TCT	GCC	ATG	CCG	GCC	TTC	ccc	TCT
136.	FGS136	CCCCCCC		TCT	ATG	ccc	GCC	****	GCC	TCT	GCT
137.	FGS137	CGCGCGC	ACATG	101	CCC	GCC	TTC	-GCC	TCT	GCT	TTC
138.	FGS138	cccccc	700110		GCC	TTC	GCC	TET	GCT	TTC	CAG
139.	FGS139	CGCGCGC	ACATG	TCT				GIT	TTC	CAG	·ccc
140.	FGS140	CGCGCGC	ACATG	TCT	TTC	GCC TCT	TCT GCT	777	CAG	·ccc	CGG
141.	FGS141	CGCGCGC	ACATG	TCT	GCC		11C	CAG	CCC	CGG	GCA
142.	FGS142	cccccc	ACATG	TCT	TCT	CCI	CAG	CAG	ccc	GCA	GGA
143.	FGS143	CGCGCGC	ACATG	TCT	GCT	TTC		csc	GCA	GGA	GGG
144.	P35144	cececec	ACATG	TCT	TTC	CAG	CGC		GGA	GGG	GTC
145.	FGS145	cececec	ACATG	TCT	CAG	ccc	CGG	SCA			
146.	FGS146	cacacac	ACATG	TCT	ccc	CCC	GCA	GGA	GGG	GTC	CTG
147.	FGS147	cececec	ACATG	TCT	CGG ·	GCA	GGA	655	GIC.	CTG	
148.	FGS148	CCCCCCC	acatg	TCT	GCA	GGA	GGG	OTC.	CIG	GTT	GCT
149.	PGS149	cececec	ACATG	TCT	GGA	GGG	GTC	·CTG	GII	GCT	AGC
150.	FGS150	CGCGCGC	ACATG	TCT	GGG	GTC	CIG	011	GCT	AGC	CAT
151.	FGS151	CGCGCGC	ACATG	TCT	<b>GIC</b>	i cre	GTT	GCT	AGC	-CAT	CTG
152.	PGS152	CGCGCGC	ACATG	TCT	CTC	GTT	GCT	ASC	CAT	CTG	CAG
153.	FGS153	CGCGCGC	ACATG	TCT	GTT	CCT	AGC	CAT	CTG	CAG	AGC
154.	PGS154	cececec	ACATG	TCT	GCT	AGC	CAT	223	CAG	AGC	130
255.	FGS155	cececec	ACATG	TCT	AGC	CAT	CTG	CXC	AGC	. TTC	·CTG
114	FGS156	2323232	ACATG	TCT	CAT	CTG	CAG	ASC	11.0	CTG	GAG
157.	FGS15"	csecce	ACATG	TCT	CTG	CAG	AGC	772	CIG	GAG	GTG
158.	PGS158	cececec	ACATG	TCT	CAG	AGC	TTC	CTG	GAG	CTC	TCG
159.	FG5159	cececec	ACATG	TCT	AGC	TTC	CTG	SAS	GTG	TCG	TAC
160.	FGS16¢	cacacac	ACATG	TCT	TTC	CTS	GAG	. 022	1700	TAC	-ccc
161.	FGS161	2000000	ACATO	201	·CTG	GAG	GTG	TES	TAC	*CGC	GTT
162.		355555	YENGS.	121 111	JAG	GTG	TCG	TAT	605	GTT	CTA
	F35161	1300000			222	TCG	TAC	525	STT	CTA	csc
163.	F351 63		ACATO							E23	CYC
154.	PG2164	cacacac	******	TCT	TCG .		. ccc	377	CTA		
165.	FGS165	CECECEC	ACATG	TCT	TAC	csc	.011	STA	CCC	EAS	CLI

166.	PGS166	cccacac	ACATG	202	CGC	GTT	CTA	CGC	CAC	<b>-CTI</b>	GCG
167.	PGS167	CGCGCGC	ACATG	TCT	CII	CTA	CGC	_CAC	c <del></del>	CCC	CAG
168.	PGS168	CECECEC	ACATG	TCT	CTA	-CCC	CAC	CII	CCC	CAG	·ccc
169.	FGS169A	CGCGCGC	ACATG	TCT	CCC	CAC	-Cil	CCC	CAG	occ	CY . C
170.	FGS170A	CGCGCCC	ACATG	TCT	CXC	·CTT	GCG	CAG	CCC	CY.C	ATG
171.	FGS171A	cececec	ACATG	TCT	CTT	GCG	CAG	ccc	.CY.C	ATG	CCT
172.	FGS172A	CCCCCCC	ACATG	TCT	GCG	CAG	∍CCC	·ey.c	ATC	CCT	ACA
173.	FGS173A	cocococ .	ACATG	TCT	CAG	CCC	CY.C	ATG	GCT	ACA	CCY .
174.	FGS174A	CGCGCGC	ACATG	TCT	CCC	Gy ⋅C	ATC	GCT	ACA	CCY	TTG
175.	FCS169B	CGCGCGC	ACATG	TCT	CCC	CAC	CIT	CCC	CAG	ccc	A CT
176.	FGS170B	CCCCCCC	ACATG	TOT	CAC	·CII	ece	CXC	-ccc	y.CI	AGT
177.	FGS1718	CGCGCGC	ACATG	TCT	CTT	CCC	CAG	ccc	Y.CL	AGT	CAT
178.	FGS172B	CCCCCCC	ACATG	TCT	GCG	CAG	ccc	Y.CL	AGT	CAT	CCA
179.	FGS173B	CCCCCCC	ACATG	TCT	CAG	ccc	A.CT	AGT	CAT	CCA	CCT
180.	FGS174B	cacacac	ACATG	TCT	CCC	Y.CL	AGT	CAT	CCA	CCT	ATG
181.	PGS169C	CGCGCGC	ACATG	TCT	CCC	CAC	CII	GCG	CXG	,000	GGC .
182.	PGS170C	CGCGCGC	ACATG	TCT	CAC	CII	ece	CAG	-000	GGC	GGC
183.	PGS171C	cececec	ACATG	TCT	CTT	GCG	CAG	CCC	GGC	GCC	GGC
184.	FGS172C	CGCGCGC	ACATG	TCT	GCG	CAG	-CCC	GGC	GGC	GGC	TCT
185.	FGS173C	CGCGCGC	ACATG	TCT	CAG	CCC	GGC	GGC	GGC	TCT	GY,C
186.	FG5174C	CGCGCGC	ACATG	TCT	CCC	GGC	CCC	GGC	TCT	GY . C	ATG
187.	RGSOA	TATATAT	GCGGCCGC	AGC	CAT	GTC	GCC	CIG	·CGC	AAG	
188.	RGSOB	TATATAT	<b>GCGGCCGC</b>	AGC	CAT	GTC	ACG	CGT	ACG	ATT	
189.	RGSOC	TATATAT	GCGGCCGC	AGC	CAT	CIC	AGA	·CCC	GCC	GCC	
190.	RGSLA	TATATAT	GCGGCCGC	TCT	AGC	CAT	CLC	GGG	CIG	CGC	
191.	RGS1B	TATATAT	GCGGCCGC	TGT	AGC	CAT	GTC	ACG	CCT	ACG	
192.	RGS1C	TATATAT	ececcec	TCT	AGC	CAT	GTC	AGA	GCC.	GCC	
193.	RGS2A	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	GGG'	CTG	
194.	RGS2B	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	ACG	-CGT	
195.	RGS2C	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	GTC	-XGX	GCC	
196.	RGS3A	TATATAT	CCCCCCCC	CAA	TGG	TGT	AGC	CAT	CTC	GGG	
197.	RGS3B	TATATAT	GCGGCCGC	CAA	TGG	TGT	AGC	CAT	GTC	ACG AGA	
198.	RGS3C	TATATAT	GCGGCCCC	CAA	TGG	TGT	AGC	CAT	GTC CAT	GTC	
199.	RGS4	TATATAT	GCGGCCGC	GCC	CAA	TGG	TGT	AGC	AGC	CAT	
200.	RGS5	TATATAT	CCCCCCC	AGG	GCC	CAA -	TGG	TGT	TGT	AGC	
201.	RGS6	TATATAT	<u>ececcec</u>	CCC	AGG	GCC	CAA	TGG	TGG	TGT	
202.	RGS7	TATATAT	GCGGCCGC	CCT	GGC	AGG	GCC	CXX	CAA	TGG	
203.	RGS8	TATATAT	ecececec	GGA	GCT	GGC	AGG	AGG	GCC	CAA	
204.	RGS9	TATATAT	GCGGCCGC	CAG	GGA	GCT		GGC	AGG	GCC	
205.	RG\$10	TATATAT	eceecec	GGG	CAG	GGA	GCT	GCT	GGC	AGG	
206.	RGS11	TATATAT	GCGGCCGC	CIG	CTG	GGG	CAG	GGA	GCT	GGC	
207.	RGS12	TATATAT	GCGGCCGC	CCT	CT	CTG	GGG	CAG	GGA	GCT	
208.	RGS13	TATATAT	GCGGCCGC	GAA	GAA	CIT	CTG	GGG	CAG	GGA	
209.	RGS14	TATATAT	GCGGCCGC	CAG GAG	CAG	GAA	GCT	CTG	GGG	CAG	
210.	RGS15	TATATAT	GCGGCCGC	CII	GAG	CAG	GAA	GCT	CTG	GGG	
211.	RGS16	TATATAT	GCGGCCGC	AGA	CTT	GAG	CAG	GAA	GCT	CTG	
212.	RGS17	TATATAT	GCGGCCGC	TAA	AGA	CIT	GAG	CAG	GAA	GCT.	
213.	RG518 RG519	TATATAT TATATAT	GCGGCCGC	CTC	TAA	AGA	CTT	GAG	CAG	GAA	
214. 215.	RGS19	TATATAT	GCGGCCGC	TTG	CTC	TAA	- AGA	CTT	GAG	·CAG	
216.	RGS21	TATATAT	GCGCCCCC	CAC	TIG	CTC	TAA	AGA	-CII	GAG	
217.	RGS22	TATATAT	GCGGCCGC	CCT	CAC	TTG	CTC	TAA	AGA	-CTT	
218.	RGS23	TATATAT	GCGGCCGC	CIT	CCT	CAC	TIG	CTC	TAA	AGA	
219.	RGS24	TATATAT	GCGGCCGC	GAT	CTT	CCT	CAC	TIC	CIC	TAA	
220.	RGS25	TATATAT	GCGGCCGC	CTG	GAT	CII	CCT	CAC	776	CIC	
221.	RGS26	TATATAT	GCGGCCGC	GCC	CTG	GAT	CTT	CCL	CYC	TTG	
222.	RG\$27	TATATAT	GCGGCCGC	ATC	GCC	CTG	GAT	CII	CCI	CYC	
223.	RGS28	TATATAT	GCGGCCGC	GCC	ATC	GCC	CIG	GAT	CTT	CCT	
224.	RGS29	TATATAT	GCGGCCGC	TGC	GCC	ATC	GCC	CTG	GAT	-CTT	
225.	RGS30	TATATAT	GCGGCCGC	CCC	TGC	GCC	ATC	ccc	CTG	GAT	
226.	RGS31	TATATAT	GCGGCCGC	GAG	CGC	TGC	GCC	ATC	GCC	CIC	
227.	RGS32	TATATAT	CCCCCCCC	CTG	GAG	CGC	TGC	GCC	ATC	-ecc	
228.	RGS33	TATATAT	GCGGCCGC	CIC	CTG	GAG	ccc	TGC	GCC	ATC	
229.	RGS34	TATATAT	GCGGCCGC	CTT	CIC	CIG	GAG	222	TGC	GCC TGC	
230.	RGS35	TATATAT	GCGGCCGC	CAG	CTT	CTC	CTG	GAG	-CGC GAG	CGC	
231.	RGS36	TATATAT	CCCCCCCC	ycx	CAG	CIT	CTC			GAG	
232.	PGS37	TATATAT	GCGCCCC	GGC	ACA	CAG	CAG	ciri	CTC CTG	CTG	
233.	RGS38	TATATAT	GCGGCCGC	GGT	OGC	YCY	ACA	CAG	CTT	CTC	
234.	RGS39	TATATAT	GCGGCCGC	GTA	GGT	GGC	GGC	ACA	CAG	CTT	
235.	RGS40	TATATAT	GCGGCCGC	CII	GTA		GGT	GGC ·	ACA	CAG	
236.	RGS41	TATATAT	GCGGCCGC	CAG	CTT	CTA CTT	GTA	GGT	GGC	ACA	
237.	R3542	TATATAT	GCGGCCGC	GCA	CAG	CAG	CTT	GTA	GGT	GGC	
138:	F2541	TATATAT	GCGGCCGC	CTG	GCX	GCA	CAG	CIT	GTA	GGT	
239.	RGS44	TATATAT	GCGGCCGC	GGG	GTG	GTG	GCY	CAG	. C11	GTA	
240.	RGS45	TATATAT	GCGGCCGC	CIC	GGG	GGG	GTG	GCA	CAS	CTT	
241.	RGS46	TATATAT	GCGGCCGC	CTC	CIC	CIC	GCG	CTG	GCA	CAG	
242.	RG547	TATATAT	ecoeccec	CAG	CAG	CTC	CIC	GGG	GTG	GCA	
243.		TATATAT	GCGGCCGC	CYC	CAC	CAG	CIC	CTC	GGG	CTC	
244.	RGS49	TATATAT	GCGGCCGC	GAG	CAG	CAC	CAG	CTC	<b>C10</b>	GGG	
245.	RGS50	TATATAT	GCGGCCGC	TCC	GAG	CAG	CAC	CAG	CTC	CTC	
246.	P.G551	TATATAT	GCGGCCGC	GTG	TCC	GAG	CAG	CAC	CAG	CTC	
247. 248.	RGS51	TATATAT TATATAT	GCGGCCGC	AGA	GTG	100	GAG	CAG	CAC	CAG	
248.	1.3553 RGS54	TATATAT	GCGGCCGC	CAG	AGA	GTG	TCC	GAG	ξλG	CAE	
250.	RGS55	TATATAT	GCGGCCGC		CAG	AGA	GTG	TCC	GAS	- CAD	
251.	RG555	TATATAT	GCGGCCGC	GAT	GCC	CAG	AGA	GTG	722	GAG	

252.	PGS57	TATATAT	CCGCCCCC	GGG	GAT	GCC	-CAG	AGA	ese	TCC
253.	RGS58	TATATAT	GCGGCCGC	CCA	GCG	SAT	CCC	CAG	AGA	CTC
254.	RGS59	TATATAT	GCGGCCGC	AGC	CCA	GGG	GAT	·GCC	-CAG	λGλ
		TATATAT	GCGGCCGC	GGG	AGC	CCA	GGG	GAT	GCC	CAG
255.	RGS60		GCGGCCGC	CAG	GGG	AGC	CCA	GGG	GAT	GCC
256.	RGS61	TATATAT				GGG	AGC	CCY	GGG	GAT
257.	RG562	TATATAT	CCCCCCC	CCT	CAG					
258.	RGS63	TATATAT	CCCCCCCC	GGA	GCT	CAG	GGG	AGC	CCY	CCC
259.	RGS64	TATATAT	GCGGCCGC	GCA	GGA	CCL	CAG	GGG	AGC	CCA
260.	RGS65	TATATAT	GCGGCCGC	GGG	GCA	GGA	GCT	CAG	GGG	AGC
		TATATAT	GCGGCCGC	GCT	GGG	GCA	GGA	GCT	CAG	GGG
261.	RGS66			CIG	GCT	GGG	GCA	GGA	GCT	-CAG
262.	RGS67	TATATAT	GCGGCCGC			CCT	GGG	GCA	GGA	GCT
263.	RGS68	TATATAT	GCGGCCGC	GGC	CIG					
264.	RGS69	TATATAT	CCGCCCCC	CAG	GGC	CTG	GCT	CCC	GCA	GGA
265.	RGS70	TATATAT	GCGGCCGC	CTG	CAG	GGC	CIG	GCT	CCC	GCA
266.	RGS71	TATATAT	GCGGCCGC	CAG	CLC	CAG	GGC	CTG	CCT	GGG
267.	RGS72	TATATAT	GCGGCCGC	TGC	CAG	CTG	CAG	GGC	-C12	·CCI
	RGS73	TATATAT	GCGGCCGC	GCC	TGC	CAG	CTG	CAG	GGC .	CTG
268.			GCGGCCGC	GCA	CCC	TGC	CAG	CTG	CAG	GGC
269.	RGS74	TATATAT			GCA	GCC	TGC	CAG	CTG	CAG
270.	RGS75	TATATAT	ccccccc	CAA					CAG	CIG
271.	RGS76	TATATAT	CCCCCCCC.	GCT	CAA	GCA	GCC	TGC		
272.	RGS77	TATATAT	CCCCCCCC	CIT	GCT	CAA	GCA	GCC	TGC	CAG
273.	RGS78	TATATAT	CCCCCCCC	GAG	GTT	GCT	CAA	GCA	GCC	TGC
274.	RGS79	TATATAT	GCGGCCGC	ATG	GAG	CTT	GCT	CAA	GCA	GCC
	RGS80	TATATAT	GCGGCCGC	GCT	ATG	GAG	GTT	CCT	CAA	GCA
.275.		TATATAT	GCGGCCGC	GCC	GCT	ATG	GAG	GTT	GCT	CAA
276.	RGS81		CCCCCCCC	AAG	GCC	GCT	ATG	GAG	GTT	GCT
277.	RGS82	TATATAT			-	GCC	GCT	ATG	GAG	GTT
278.	RGS83	TATATAT	ececcec	GAA	AAG					
279.	RGS84	TATATAT	ececcec .	GAG	GAA	AAG	GCC	GCT	ATG	GAG
280.	RG585	TATATAT	CCCCCCCC	GTA	GAG	GAA	AAG	GCC	GCT	ATG
281.	RGS86	TATATAT	GCGGCCGC	CIG	GTA	GAG	GAA	AAG	GCC	GCT
282.	RGS87	TATATAT	GCGGCCGC	CCC	CTG	GTA	GAG	GAA	AAG	GCC
	RGS88	TATATAT	GCGGCCGC	GAG	ccc	CTG	GTA	GAG	GAA	AAG
283.			GCGGCCGC	CAG	GAG	CCC	CTG	GTA .	GAG	GAA
284.	RG589	TATATAT		CTG	CAG	GAG	CCC	CTG	GTA	GAG
285.	RGS90	TATATAT	GCGGCCGC					CCC	CTG	GTA
286.	RGS91	TATATAT	ececcec	GGC	CTG	CAG	GAG			
287.	RGS92	TATATAT	CCCCCCC	CAG	GGC	CTG	CAG	GAG	CCC	CTG
288.	RGS93	TATATAT	GCGGCCGC	TTC	CAG	GCC	CTG	CAG	GAG	-CCC
289.	RGS94	TATATAT	GCGGCCGC	CCC	TTC	CAG	GGC	CTG	CAG	GAG
290.	RGS95	TATATAT	GCGGCCGC	TAT	CCC	TTC	CAG	GGC	CTG	CAG
291.	RGS96	TATATAT	GCGGCCGC	GGA	TAT	CCC	TTC	CAG	GGC	CIG
		TATATAT	GCGGCCGC	GGG	GGA	TAT	ccc	TTC	CAG	GGC
292.	RG597		GCGGCCGC	CIC	GGG	GGA	TAT	CCC	TTC	CAG
293.	RGS98	TATATAT				GGG	GGA	TAT	CCC	TTC
294.	RGS99	TATATAT	ecceccec.	CAA	CTC					-000
295.	RGS100	TATATAT	<b>GCGGCCGC</b>	ACC	CYY	CIC	GGG	GGA	TAT	
296.	RGS101	TATATAT	GCGGCCGC	GGG	ACC	CAA	CTC	GGG	GGA	TAT
297.	RGS102	TATATAT	GCGGCCGC ·	GGT	GGG	ACC	CAA	CTC	GGG	GGA
298.	RGS103	TATATAT	CCCCCCC	CAA	GGT	GGG	ACC	CAA	CTC	GGG
299.	RGS104	TATATAT	GCGGCCGC	GTC	CAA	GGT	GGG	ACC	CAA	CTC
		TATATAT	GCGGCCGC	TGT	GTC	CAA	GCT	GGG	ACC	CAA
300.	RGS105		GCGGCCGC	CAG	TGT	GTC	CAA	GGT	GGG	ACC
301.	RGS106	TATATAT				TGT	GTC	CAA	GGT	GGG
	RG5107	TATATAT	GCGGCCGC	CTG	CAG				CAA	GGT
303.	RGS108	TATATAT	GCGGCCGC	CAG	CTG	CAG	TCT	GTC		
304.	RGS109	TATATAT	GCGGCCGC	GTC	CXG	CTG	CAG	TCT	·GT=	CAA
305.	RGS110	TATATAT	GCGGCCGC	GAC	CTC	CAG	CTG	CAG	TGT	GTC
306.	RGS111	TATATAT	GCGGCCGC	GGC	GAC	GTC	CAG	CTG	CAG	ici
307.	RGS112	TATATAT	GCGGCCGC	GTC	GGC	GAC	GTC	CAG	CIC	-CAG
308.	RGS113	TATATAT	GCGGCCGC	AAA	GTC	GGC	GAC	GTC	CAG	CTG
		TATATAT	GCGGCCGC	GGC	AAA	GTC	GGC	GAC	GTC	CAG
309.	RGS114				GGC ·	AAA	GTC	GGC	GAC	GTC
310.	RGS115	TATATAT	GCGGCCGC	GGT		GGC	AAA	GTC	GGC	GAC
311.	RG5116	TATATAT	GCGGCCGC	GGT	CGT		GGC		GTC	GGC
312.	RGS117	TATATAT	GCGGCCGC	GAT	GGT	GGT		XXX		GTC
313.	RGS118	TATATAT	CCCCCCCC	CCA	GAT	GGT	GGT	GGC	XXX	
314.	RGS119	TATATAT	CCCCCCCC	CTG	CCY	GAT	GGT	GGT	GGC	AAA
315.	RGS120	TATATAT	GCGGCCGC	CTG	CIG	CCY	GAT	GGT	GGT	GGC
316.	RGS121	TATATAT	<b>CCCCCCC</b>	CAT	CTG	CTG	CCA	GAT	GCT	GGT
317.	RGS122	TATATAT	GCGGCCGC	TTC	CAT	CIG	CTG	CCA	GAT	GGT
318.	RG5123	TATATAT	GCGGCCGC	TTC	TTC	CAT	CTG	CTG	-CCA	GAT
		TATATAT	GCGGCCGC	CAG	TTC	TTC	CAT	CTG	CTG	CCA
319.	P.GS124		GCGGCCGC	100	CAG	TTC	TTC	CAT	CTG	CTG
320.	.RGS125	TATATAT			TCC	CAG	110	770	CAT	CIG
321.	RGS126	TATATAT	GCGGCCGC	CAT						
322.	RGS127	TATATAT	GCGGCCGC	GGC	CAT	TCC	CAG	TTC	TTC	CAT
323.	RGS128	TATATAT	GCGGCCGC	AGG	GGC	CAT	TCC	CYC	TTC	TTC
324.	RGS129	TATATAT	GCGGCCGC	GGC	AGG	GGC	CAT	TCC	CAG	TIC
325.	R55130	TATATAT	GCGGCCGC	CAG	GGC	AGG	GGC	CAT	TCC	'CAG
326.		TATATAT	GCGGCCGC	CIG	CAG	GGC	AGG	GGC	CAT	· TCC
		TATATAT	CCCCCCC	GGG	CTG	CAG	GGC	AGG	GGE	CAT
327.				GGT	GGG	CTG	CAG	GGC	AGG	GGC
328.		TATATAT	GCGGCCGC							
329.		TATATAT	ececcec	CTG	GCT	GGG	CIG	CAG	GGC	AGG
330.	RGS135	TATATAT	GCGGCCGC	YCC	CTG	GGT	GGG	CTG	CAG	GGC
331.		TATATAT	CCCCCCCC	GGC	ACC	CIG	GGT	GGG	-620	್ಲಿಯ
332.		TATATAT	GCGGCCGC	· CAT	GGC	ACC	CTG	CCT	GGG	CTG
333.		TATATAT	GCGGCCGC	CGG	CAT	GGC	ACC	CTG	GCT	GGG
			GCGGCCGC	GGC	CGG	CAT	GGC	ACC	CTG	GGT
334.		TATATAT			CCC	-662			ACE	cre.
335.		TATATAT	GCGGCCGC	GAA			CAT	GGC		
3:6.		TATATAT	ecaecaec	GSC	GAA	GGC	CGG	-CAT	950	YCC
337.	RGS142	TATATAT	GCGGCCGC	AGA	CCC	GAA	GGC	CGG	CAT	300

338.	RGS143	TATATAT	GCGGCCGC	AGC	AGA	GGC	GAA	GGC	ccc	CAT	
339.	RGS144	TATATAT	GCGGCCGC	GAA	AGC	AGA	GGC	GAA	CCC	CGG	
340.	RGS145	TATATAT	CCCCCCCC	CTG	GAA	AGC	AGA	~ GGC	GAA	GGC	
	RGS146	TATATAT	CCCCCCC	GCG	CTG	GAA	AGC	AGA	GCC	GAA	
341.	RGS147	TATATAT	GCGCCCCC	CCG	GCG	CTG	- GAA	AGC	AGA	CCC	
342.	RGS148	TATATAT	CCGCCCCC	TGC	CCG	GCG	·CTG	GAA	AGC	AGA	
343.	RGS149	TATATAT	GCGGCCGC	TCC	TGC	CCG	GCG	CIG	GAA	AGC	
344.	RGS150	TATATAT	GCGGCCGC	CCC	TCC	TGC	CCC	GCG	CTG	GAA	
345.		TATATAT	GCGGCCGC	GAC	CCC	TCC	TGC	ccc	GCG	CTG	•
346.	RGS151	TATATAT	GCGGCCGC	CAG	GAC	ccc	TCC	TGC	CCG	GCG	
347.	RGS152	TATATAT	GCGGCCGC	AAC	CAG	GAC	CCC	TCC	TGC	CCG	
348.	RGS153	TATATAT	GCGGCCGC	AGC	AAC	CAG	GAC	-000	TCC	TGC	
349.	RGS154	TATATAT	GCGGCCGC	GCT	AGC	AAC	CAG	GAC	CCC	TCC	
350.	RG5155		GCGGCCGC	ATG	GCT	AGC	AAC	CAG	GAC	CCC	
351.	RGS156	TATATAT TATATAT	GCGGCCGC	CAG	ATG	GCT	AGC	AAC	CAG	GAC	-
352.	RGS157	TATATAT	GCGGCCGC	CTG	CAG	ATG	CCT	AGC	AAC	CAG	
353.	RGS158	TATATAT	GCGGCCGC	GCT	CTG	CAG	ATG	CCT	AGC	AAC	
354.	RGS159		GCGGCCGC	GAA	GCT	CTG	CAG	ATG	CCT	AGC	
355.	RGS160	TATATAT	GCGGCCGC	CAG	GAA	GCT	CTG	CAG	ATG	GCT	
356.	RGS161	TATATAT	GCGGCCGC	CTC	CAG	GAA	GCT	CTG	CAG	ATG	
357.	RGS162	TATATAT	GCGGCCGC	CAC	CTC	CAG	GAA	GCT	CIG	CAG	
358.	RGS163	TATATAT	GCGGCCGC	CGA	CAC	CTC	CAG	GAA	GCT	-CTG	
359.	RGS164	TATATAT	GCGGCCGC	GTA	CGA	CAC	CTC	CAG	GAA	GCT	
360.	RGS165	TATATAT	GCGGCCGC	GCG	GTA	CGA	CAC	CTC	CAG	GAA	
361.	RGS166	TATATAT	GCGGCCGC	AAC	GCG	GTA	CGA	CAC	CTC	CAG	
362.	RGS167	TATATAT		TAG	· AAC	GCG	GTA	CGA	CAC	CTC	•
363.	RGS168	TATATAT	GCGGCCGC	GCG	TAG	AAC	GCG	GTA	CGA	CAC	
364.	RGS169	TATATAT	GCGGCCGC	GTG	GCG	TAG	AAC	GCG	CTA	CGA	
365.	RGS170	TATATAT	GCGGCCGC	AAG	GTG	GCG	TAG	AAC	GCG	GTA	
366.	RG5171	TATATAT	GCGGCCGC	CCC	AAG	GTG	GCG	TAG	AAC	GCG	
367.	RGS172	TATATAT	GCGGCCGC	CTG	CGC	AAG	GTG	GCG	TÀG	AAC	GCG
368.	PGS173	TATATAT	CCGCCCCC	C10		- AAU	3.0				

5

10

15

20

#### **CLAIMS**

#### What is claimed is:

1. A method for making a biologically-active circularly-permuted protein of the formula C¹-L¹-N¹, derived from a parent protein of the formula N¹-C¹, wherein

C<sup>1</sup> is comprised of a segment derived from the carboxy portion of said parent protein;

 $\mathbf{N}^{1}$  is comprised of a segment derived from the amino terminal portion of said parent protein; and

 $L^1$  is a chemical bond or a linker, linking  $C^1$  to the amino terminus of  $L^1$  and carboxy terminus of  $L^1$  to the amino terminus of  $N^1$ ;

comprising the steps of:

- (a) making a series of circularly-permuted genes;
- (b) inserting said circularly-permuted genes into a display vector;
- (c) expressing said circularly-permuted genes such that the proteins encoded by said genes are presented on the surface of the display vector;
- (d) generating a library of display vectors presenting the expressed circularly permuted protein;
- (e) affinity-selecting the presenting display vectors with a target protein that can bind a biologically-active circularly-permuted protein;
  - (f) isolating and analyzing clones of selected display vectors to identify the presented circularly-permuted protein.
- 2. The method of claim 1 wherein the method of making a series of circularly-permuted genes is selected from the group consisting of making a tandemly-repeated intermediate, total synthesis of a synthetic gene, assembly of a gene from synthetic oligonucleotides, DNA amplification, and limited digestion of a circular

intermediate.

- The method of claim 1 wherein said display vector is selected from the group consisting of bacteriophage display vectors, bacteria, and baculovirus vectors.
  - 4. The method of claim 3, wherein said presentation vector is a bacteriophage.
    - The method of claim 4, wherein said presentation vector is bacteriophage M13.
      - 6. The method of claim 5, wherein said presentation vector is a bacteriophage M13 gene III vector.
- 7. The method of claim 1 wherein said method of making a series of circularly permuted genes is a method of making a tandem repeat intermediate.
  - The method of claim 7 wherein said circularly-permuted genes are amplified from the repeat by gene amplification.
- 9. The method of claim 1 wherein said method of affinity selection comprises the steps consisting of:
  - (a) binding said presentation display vectors to a target protein;
  - (b) eluting said display vectors;
  - (c) amplifying said display vectors; and
  - (d) biopanning a pool of said amplified display vectors.
  - 10. The method of claim 1 wherein  $L^1$  is a linear peptide linker.
  - 11. The method of claim 1 wherein said the DNA sequence encoding said linker L<sup>1</sup> is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 368.
  - 12. The method of claim 1 wherein the length of the  $C^1$  in said permutein is shorter than the length of  $C^1$  in said parent protein.

10

5

15

20

25

. 5

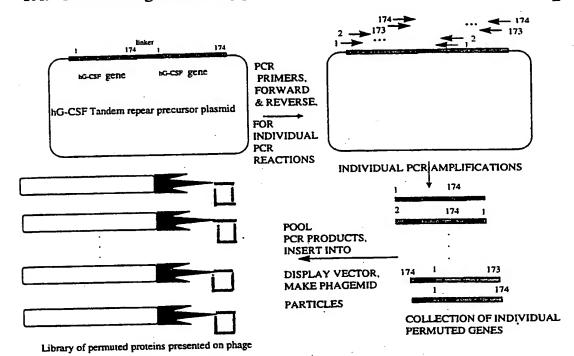
- 13. The method of claim 1 wherein the length of the  $N^1$  in said permutein is shorter than the length of  $N^1$  in said parent protein.
- 14. A circularly-permuted protein prepared by the method of claim 1.
  - 15. A circularly-permuted protein of claim 14 comprising the G-CSF receptor agonist domain of a species of mylepoietin (MPO).

1/4

### **FIGURES**

### Figure 1A

# 1A. Constructing a scanning permutagenesis display library



PCT/US99/20891

## 2/4 Figure 1B

# 1B. Screening a display library

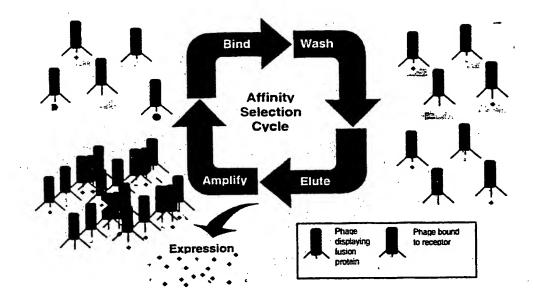


Figure 2

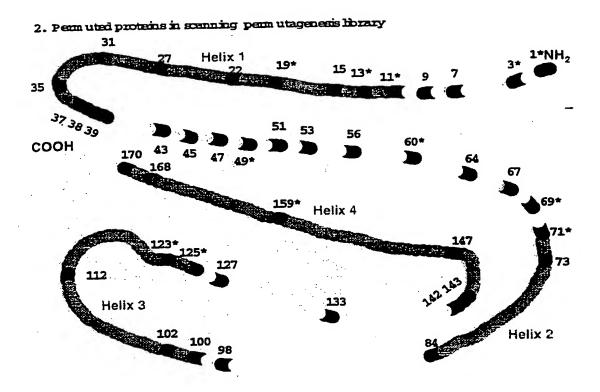
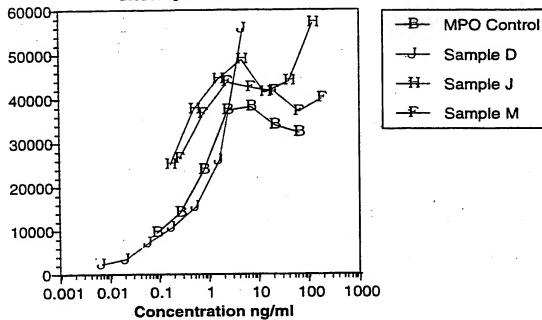


Figure 3





DIEDOCIO ANO OCTODENTI

### SEQLIST.TXT

### SEQUENCE LISTING

CITO Bee, Scopius.	
and the parameters but	
<120> Method of Producing Permuteins by Scanning Permutagenesis	
<130> C_3041	
<140> 60/	
<141> 1998-09-25	
<160> 368	
<170> FastSEQ for Windows Version 3.0	
<210> 1	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 1 cgcgcgcaca tgtctacacc attgggccct gccagc	36
<210> 2	
<211> 36	
<212> DNA	•
<213> Synthetic	
<400> 2	3.6
cgegegeaca tgtetecatt gggeeetgee agetee	36
<210> 3	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 3	2.6
cgcgcgcaca tgtctttggg ccctgccagc tccctg	36
<210> 4	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 4	36
cgcgcgcaca tgtctggccc tgccagctcc ctgccc	
<210> 5	
<211> 3€	
<212> DNA	
<213> Synthetic	
<400> 5	36
cgcgcgcaca tgtctcctgc cagctccctg ccccag	
<210> 6	
<211> 36	
<212> DNA	
<213> Synthetic	
<4GC> 6	*
cgcgcgcaca tgtctgccag ctccctgccc cagagc	3

<210> 7				•	
<211> 36			•		
<212> DNA					
<213> Synthetic					
(213) Synchecic					
<400> 7					2.6
cgcgcgcaca tgtctagctc	cctgccccag	agcttc			3.6
<210> 8			,		
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 8					
cgcgcgcaca tgtcttccct	gccccagagc	ttcctg			36
-				•	
<210> 9					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 9					•
cgcgcgcaca tgtctctgcc	ccagagcttc	ctgctc		•	36
		•			
<210> 10					
<211> 36					
<212> DNA					
<213> Synthetic					
. <400> 10					
cgcgcgcaca tgtctcccca	gagetteetg	ctcaag			36
<210> 11					
<211> 36					
<212> DNA					
<213> Synthetic			•		
<213> Synthetic					
				•	
<400> 11					´36
cgcgcgcaca tgtctcagag	cttectgete	aagtct			30
<210> 12					
<211> 36					
<212> DNA					
<213> Synthetic				•	
12137 Synthetic					
-400- 32	•	-			
<400> 12					36
cgcgcgcaca tgtctagctt	cetgeteaag	tettea			
•					
<210> 13			•		
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 13					•
		******			36
cgcgcgcaca tgtctttcct	geteaagtet	. ccayay			
			į	•	
<230> 14					
<211> 36			•		
<212> DNA		•			
<213> Synthetic	•				
<400> 14			•		
		02555			36
cgcgcgcaca tgtctctgct	. caagtettt	yaycaa	•		, 30
<210> 15					
<211> 36			•		
<212> DNA					

	_	
<400> 15	•	36
cgcgcgcaca tgtctctcaa gtctttagag caagtg	_	, 0
<210> 16		
<211> 36		
<212> DNA		
<213> Synthetic		
1000		
<400> 16		
cgcgcgcaca tgtctaagtc tttagagcaa gtgagg		36
egegegeaca egectaaget transmission session		
.030- 17		
<210> 17	•	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 17		36
cgcgcgcaca tgtcttcttt agagcaagtg aggaag	•	30
<210> 18		
<211> 36		
<212> DNA	•	
<213> Synthetic		
4215	•	
<400> 18		
cgcgcgcaca tgtctttaga gcaagtgagg aagatc		3′6
CGCGCGCACA CGCCCCAGA GCAAGAGAGAGAGAGAGAG		
.010- 10		
<210> 19		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 19		3'6
cgcgcgcaca tgtctgagca agtgaggaag atccag		
	• .	
<210> 20		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 20		
cgcgcgcaca tgtctcaagt gaggaagatc cagggc	•	36
<210> 21		
<211> 36		
<212> DNA		
<213> Synthetic		
(213) Synthetia		
-400> 31		
<400> 21 cgcgcgcaca tgtctgtgag gaagatccag ggcgat	•	3.6
cgcgcgcaca tgtctgtgag gaagattaag ggogat		
<210> 22		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 22		2.0
cgcgcgcaca tgtctaggaa gatccagggc gatggc		36
<210> 23		
<211> 36		
<212> DNA		
<213> Synthetic		
-pag- ojuduoteo		
<400> 23		
		36
cgcgcgcaca tgtctaagat ccagggcgat ggcgca		

<210> 24

<400> 32

<i>i</i>						
· <211> 36				•		
<212> DNA						
<213> Synthetic						
<400> 24						
cgcgcgcaca tgtctatcca	gggcgatggc	gcagcg				36
-3-3-3						
<210> 25						
<211> 36	•					
<212> DNA						
<213> Synthetic						
<400> 25						
cgcgcgcaca tgtctcaggg	cgatggcgca	gcgctc				36
<210> 26				•		
<211> 36						
<212> DNA						
<213> Synthetic						
<213> Synthetic						
-400- 25		•				
<400> 26						26
cgcgcgcaca tgtctggcga	rggcgcagcg	ctccag			•	36
<210> 27						
<211> 36						
<212> DNA						
<213> Synthetic						
_						
<400> 27						
cgcgcgcaca tgtctgatgg	cacaacactc	caggag				36
0909090000 0900090099	-5-5-5-5	33-5.				
<210> 28						
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 28				:		
cgcgcgcaca tgtctggcgc	agcgctccag	gagaag				36
<210> 29		•				
<211> 36	•					
<212> DNA						
<213> Synthetic						
_						
<400> 29						
cgcgcgcaca tgtctgcago	getecaggag	AAGCEG				36
cycycycaca cytolycay	. 30.000.35			•		
<210> 30						
				•		
<211> 36						
<212> DNA						
<213> Synthetic	:					
<400> 30						
cgcgcgcaca tgtctgcgct	ccaggagaag	ctgtgt				36
<210> 31						
<211> 36		•	1			
<212> DNA						
<213> Synthetic	•					
Jillieux	-					
<400> 31						
		tataca	•			36
cgcgcgcaca tgtctctcca	yyayaagctg	Lytyce				30
			•			
<210> 32						
-211s 2E						
<211> 36						
<212> DNA		•				
•		•	- •			

cgcgcgcaca tgtctcagga gaagctgtgt gccacc	36
<210> 33	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 33	26
cgcgcgcaca tgtctgagaa gctgtgtgcc acctac	36
<210> 34	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 34	36
cgcgcgcaca tgtctaagct gtgtgccacc tacaag	
.010. 25	
<210> 35	
<211> 36 <212> DNA	
<213> Synthetic	
(21) Synthecat	
<400> 35	
cgcgcgcaca tgtctctgtg tgccacctac aagctg	36
Cgcgcgcaaa cgccccc	
<210> 36	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 36	36
egegegeaca tgtettgtge cacetacaag etgtge	30
nn	
<210> 37	
<211> 36 <212> DNA	
<213> Synthetic	
<213> Synthetic	
<213> Synthetic <400> 37	36
<213> Synthetic	36
<213> Synthetic <400> 37	36
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac	36
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38	36
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38  <211> 36	36
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38 <211> 36 <212> DNA <213> Synthetic	36
<213> Synthetic  <400> 37 cgcgegcaca tgtctgccac ctacaagctg tgccac  <210> 38 <211> 36 <212> DNA <213> Synthetic  <400> 38	
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38 <211> 36 <212> DNA <213> Synthetic	36
<213> Synthetic  <400> 37 cgcgegcaca tgtctgccac ctacaagctg tgccac  <210> 38  <211> 36  <212> DNA  <213> Synthetic  <400> 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc	
<213> Synthetic  <400> 37 cgcgegcaca tgtctgccac ctacaagctg tgccac  <210> 38  <211> 36  <212> DNA  <213> Synthetic  <400> 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc  <210> 39	
<213> Synthetic  <400> 37 cgcgegcaca tgtctgccac ctacaagctg tgccac  <210> 38 <211> 36 <212> DNA <213> Synthetic  <400> 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc  <210> 39 <211> 36	
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38 <211> 36 <212> DNA <213> Synthetic  <400> 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc  <210> 39 <211> 36 <212> DNA	
<213> Synthetic  <400> 37 cgcgegcaca tgtctgccac ctacaagctg tgccac  <210> 38 <211> 36 <212> DNA <213> Synthetic  <400> 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc  <210> 39 <211> 36	
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38	36
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38 <211> 36 <212> DNA <213> Synthetic  <400> 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc  <210> 39 <211> 36 <212> DNA <213> Synthetic	
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38</pre>	36
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38 <211> 36 <212> DNA <213> Synthetic  <400> 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc  <210> 39 <211> 36 <212> DNA <213> Synthetic	36
<213> Synthetic  <400> 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  <210> 38	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgegcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgegcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38</pre>	36 36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgegcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc  &lt;210&gt; 39 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 39 cgcgcgcaca tgtctaccta caagctgtgc cacccc  &lt;210&gt; JNA &lt;213&gt; Synthetic  &lt;400&gt; 39 cgcgcgcaca tgtcttacaa gctgtgccac cccgag  &lt;210&gt; 40 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 40 cgcgcgcaca tgtcttaagct gtgccacccc gaggag</pre>	36 36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38</pre>	36 36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 37 cgcgcgcaca tgtctgccac ctacaagctg tgccac  &lt;210&gt; 38 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 38 cgcgcgcaca tgtctaccta caagctgtgc cacccc  &lt;210&gt; 39 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 39 cgcgcgcaca tgtctaccta caagctgtgc cacccc  &lt;210&gt; JNA &lt;213&gt; Synthetic  &lt;400&gt; 39 cgcgcgcaca tgtcttacaa gctgtgccac cccgag  &lt;210&gt; 40 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 40 cgcgcgcaca tgtcttaagct gtgccacccc gaggag</pre>	36 36

<213> Synthetic	<u>.</u>	
<400> 41		
cgcgcgcaca tgtctctgtg ccaccccgag ga	gctg	36
<210> 42		
<211> 36		
<212> DNA		
<213> Synthetic		
4005 40		
<400> 42 cgcgcgcaca tgtcttgcca ccccgaggag ct	agta	36
cycycycaca cyccecycca ccccycygus	.33-3	
<210> 43		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 43	racto	36
cgcgcgcaca tgtctcaccc cgaggagctg gt	.gc.cg	-
<210> 44		
<211> 36	•	
<212> DNA		
<213> Synthetic		
<400> 44	·	3.6
egegegeaca tgteteeega ggagetggtg e	igete	30
<210> 45		
<211> 36		
<212> DNA		
<213> Synthetic		
-		
<400> 45		36
cgcgcgcaca tgtctgagga gctggtgctg c	ccgga	-
<210> 46		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 46		36
cgcgcgcaca tgtctgagct ggtgctgctc g	gacac	,,
<210> 47		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 47		36
cgcgcgcaca tgtctctggt gctgctcgga c	actet	30
<210> 48		
<211> 46		
<212> DNA		
<213> Synthetic	,	
	•	
<400> 48		2.5
cgcgcgcaca tgtctgtgct gctcggacac 1	ectetg	3.6
c2105 AD		
<210> 49 <211> 36		
<211> 36 <212> DNA		
<213> Synthetic		
<400> 49		
		36

<210> 50 .						
<211> 36				-		
<212> DNA						
<213> Syntheti	C					
						•
<400> 50						
		ecet.c				36
cgcgcgcaca tgtctctcg	g acactetety	ggcacc				50
	,					
<210> 51						
<211> 36						
					•	
<212> DNA						
<213> Syntheti	С					
<400> 51						
cgcgcgcaca tgtctggac	a ctctctgggc	atcccc				36
cycycycaca cytecyguo	u 000000,,,,					
<210> 52			•			
<211> 36						
<212> DNA						
<213> Syntheti	_			•		
(213) Syntheti	_					
<400> 52					•	
cgcgcgcaca tgtctcact	c tetgggcate	ccctgg				36
<210> 53	•					
<211> 36						
<212> DNA						
<213> Syntheti	.c					
<400> 53						
		tagact				36
cgcgcgcaca tgtcttctc	t gygtattet	rgggc				
<210> 54						
<211> 36						
<211> 36 <212> DNA						
<211> 36	. <b>c</b>					
<211> 36 <212> DNA <213> Syntheti	. <b>c</b>					
<211> 36 <212> DNA <213> Syntheti <400> 54						
<211> 36 <212> DNA <213> Syntheti <400> 54		getece				36
<211> 36 <212> DNA <213> Syntheti		getece				36
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg		getece				36
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55		getece				36
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36		getece				36
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55		getece				36
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA	g catecectgg	getece			*	36
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36	g catecectgg	getece			*	36
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Syntheti	g catecectgg	getece			*	36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgcctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic	gg cateceetgg					
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Syntheti	gg cateceetgg					36
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Syntheti <400> 55 cgcgcgcaca tgtctgccaca tgtctctgccaca tgtctctgccaca tgtctctgccaca tgtctctgccaca tgtctgccaca tgtctggcaca tgtctggcaca tgtctggcaca	gg cateceetgg				*	
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgcctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic	gg cateceetgg					
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgcctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgcctggcaca tgcctggcaca tgcctggcaca tgcctggcaca tgcctggcaca tgcctggcaca cgccgcgcaca tgcctggcaca cgccgcgcaca cgccgcgcaca tgcctggcaca cgccgcgcaca cgcccgcgcaca cgccgcgcaca cgccgcaca cgccgcgcaca cgccgcgcaca cgccgcgcaca cgccgcgcaca cgccgcgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgccgcaca cgccgcaca cgccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccgcaca cgcccaca cgcccgcaca cgcccgcaca cgcccaca cgccaca cgcccaca cgccaca cqccaca cqccaca cqccaca cqccaca cqccaca cqccaca cqccaca cqccaca cqccaca cqccac	gg cateceetgg					
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca cg10> 56 <211> 36 <210> 56 <211> 36	gg cateceetgg					
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcacaca tgtctggcacacacacacacacacacacacacacacacac	gg cateceetgg ie					
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca cg10> 56 <211> 36 <210> 56 <211> 36	gg cateceetgg ie					
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Syntheti <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcacaca tgtctggcacacacacacacacacacacacacacacacac	gg cateceetgg ie					
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca cg10> 56 <211> 36 <212> DNA <213> Synthetic <400> 56	gg catecectgg ic at eccetggget	ccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca cg10> 56 <211> 36 <212> DNA <213> Synthetic <400> 56	gg catecectgg ic at eccetggget	ccctg				
<211> 36 <212> DNA <213> Syntheti <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Syntheti <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcacaca tgtctggcacacacacacacacacacacacacacacacac	gg catecectgg ic at eccetggget	ccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgcctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgcctggcaca tgcctggcaca tgcctggcaca <210> 56 <211> 36 <212> DNA <213> Synthetic <400> 56 cgcgcgcaca tgcctggcaca tgcctggcaca tgcctggcaca tgcctggcaca tgcctggcaca tgcctacaca tgcctacacaca tgcctacaca tgcctacacaca tgcctacacaca tgcctacacaca tgcctacacacacacacacacacacacacacacacacaca	gg catecectgg ic at eccetggget	ccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctgcaca cgcgcgcaca tgtctatcaccacacaca tgtctatcaccacacacacacacacacacacacacacaca	gg catecectgg	ccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctaccacacacacacacacacacacacacacacacac	gg catecectgg	ccctg				36
<pre>&lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 54 cgcgcgcaca tgtctctgg &lt;210&gt; 55 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 55 cgcgcgcaca tgtctggcaca tgtctggcaca &lt;210&gt; 56 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 56 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctatcaca &lt;210&gt; 56 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 56 cgcgcgcaca tgtctatcacacacacacacacacacacacacacacacac</pre>	gg catecectgg  ic  ic  ic  cectggget	ccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctaccacacacacacacacacacacacacacacacac	gg catecectgg  ic  ic  ic  cectggget	ccctg				36
<pre>&lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 54 cgcgcgcaca tgtctctgg &lt;210&gt; 55 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 55 cgcgcgcaca tgtctggcaca tgtctggcaca &lt;210&gt; 56 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 56 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctatcaca &lt;210&gt; 56 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 56 cgcgcgcaca tgtctatcacacacacacacacacacacacacacacacac</pre>	gg catecectgg  ic  ic  ic  cectggget	ccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctacacacacacacacacacacacacacacacacaca	gg catecectgg  ic  ic  ic  cectggget	ccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgcctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgcctggcaca tgcctggcaca <210> 56 <211> 36 <212> DNA <213> Synthetic <400> 56 cgcgcgcaca tgcctggcaca <210> 56 <211> 36 <212> DNA <213> Synthetic <400> 56 cgcgcgcaca tgcctatcaca <210> 57 <211> 36 <212> DNA <213> Synthetic <400> 57 <211> 36 <212> DNA <213> Synthetic <400> 57	gg catecectgg  ic  ce etgggetece	cccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctacacacacacacacacacacacacacacacacaca	gg catecectgg  ic  ce etgggetece	cccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgcctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgcctggcaca tgcctggcaca <210> 56 <211> 36 <212> DNA <213> Synthetic <400> 56 cgcgcgcaca tgcctggcaca <210> 56 <211> 36 <212> DNA <213> Synthetic <400> 56 cgcgcgcaca tgcctatcaca <210> 57 <211> 36 <212> DNA <213> Synthetic <400> 57 <211> 36 <212> DNA <213> Synthetic <400> 57	gg catecectgg  ic  ce etgggetece	cccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgcctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgcctggcaca tgcctggcaca <210> 56 <211> 36 <212> DNA <213> Synthetic <400> 56 cgcgcgcaca tgcctggcaca <210> 56 <211> 36 <212> DNA <213> Synthetic <400> 56 cgcgcgcaca tgcctatcaca <210> 57 <211> 36 <212> DNA <213> Synthetic <400> 57 <211> 36 <212> DNA <213> Synthetic <400> 57	gg catecectgg  ic  ce etgggetece	cccctg				36
<pre>&lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti  &lt;400&gt; 54 cgcgcgcaca tgtctctgg &lt;210&gt; 55 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti  &lt;400&gt; 55 cgcgcgcaca tgtctggca &lt;210&gt; 56 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 56 cgcgcgcaca tgtctggca &lt;210&gt; 56 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 56 cgcgcgcaca tgtctatca &lt;210&gt; 57 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 57 cgcgcgcaca tgtctatca &lt;210&gt; 57 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 57 cgcgcgcaca tgtctccc &lt;210&gt; 57</pre>	gg catecectgg  ic  ce etgggetece	cccctg				36
<211> 36 <212> DNA <213> Synthetic <400> 54 cgcgcgcaca tgtctctgg <210> 55 <211> 36 <212> DNA <213> Synthetic <400> 55 cgcgcgcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcaca tgtctggcacacaca tgtctacacacacacacacacacacacacacacacacaca	gg catecectgg  ic  ce etgggetece	cccctg				36
<pre>&lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti  &lt;400&gt; 54 cgcgcgcaca tgtctctgg &lt;210&gt; 55 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti  &lt;400&gt; 55 cgcgcgcaca tgtctggca &lt;210&gt; 56 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 56 cgcgcgcaca tgtctggca &lt;210&gt; 56 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 56 cgcgcgcaca tgtctatca &lt;210&gt; 57 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 57 cgcgcgcaca tgtctatca &lt;210&gt; 57 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Syntheti &lt;400&gt; 57 cgcgcgcaca tgtctccc &lt;210&gt; 57</pre>	gg catecectgg  ic  at eccetggget  ic  cc etgggetece  ic	cccctg				36

<400> 58 cgcgcgcaca tgtcttgggc				
cacacacaca tacccagage	tcccctgagc	tcctgc	-	3.6
<210> 59				
<211> 36				
<212> DNA			•	
<213> Synthetic				
<400> 59				36
cgcgcgcaca tgtctgctcc	cctgagctcc	tgcccc		50
<210> 60				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 60				
cgcgcgcaca tgtctcccct	gagctcctgc	cccagc		36
<210> 61				
<211> 36		•		
<212> DNA				
<213> Synthetic			•	
(225) 07.10110550				
<400> 61				
cgcgcgcaca tgtctctgag	crectacec	agecag		36
egegegeaca egeceegag	Ctcccgcccc	ug00ug		
-210> 62				
<210> 62				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 62				36
cgcgcgcaca tgtctagctc	ctgccccagc	caggcc		30
<210> 63				
<211> 36				
<212> DNA				
<213> Synthetic				
			•	
<400> 63				
	ccccagecag	geeetg	·	36
<400> 63 cgcgcgcaca tgtcttcctg	ccccagccag	geeetg	·	36
	ccccagccag	gecetg		36
cgcgcgcaca tgtcttcctg	cccagccag	gecetg		36
cgcgcgcaca tgtcttcctg <210> 64 <211> 36	ccccagccag	geeetg		36
<pre>cgcgcgcaca tgtcttcctg       &lt;210&gt; 64       &lt;211&gt; 36       &lt;212&gt; DNA</pre>	٠	gecetg		36
cgcgcgcaca tgtcttcctg <210> 64 <211> 36	٠	gecetg		36
<pre>cgcgcgcaca tgtcttcctg &lt;210&gt; 64 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	٠	gecetg		36
<pre><gcgcgcaca <210="" tgtcttcctg=""> 64 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 64</gcgcgcaca></pre>				36
<pre>cgcgcgcaca tgtcttcctg &lt;210&gt; 64 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>				
<pre>cgcgcgcaca tgtcttcctg      &lt;210&gt; 64      &lt;211&gt; 36      &lt;212&gt; DNA      &lt;213&gt; Synthetic      &lt;400&gt; 64 cgcgcgcaca tgtcttgccc</pre>				
<pre><gcgcgcaca td="" tgtcttcctg<=""><td></td><td></td><td></td><td></td></gcgcgcaca></pre>				
<pre><gcgcgcaca td="" tgtcttcctg<=""><td></td><td></td><td></td><td></td></gcgcgcaca></pre>				
<pre>&lt;210&gt; 64</pre>	cagccaggco			
<pre><gcgcgcaca td="" tgtcttcctg<=""><td>cagccaggco</td><td></td><td></td><td></td></gcgcgcaca></pre>	cagccaggco			
<pre>&lt;210&gt; 64</pre>	cagccaggco			
<pre></pre>	cagccaggco	ctgcag		36
<pre>&lt;210&gt; 64</pre>	cagccaggco	ctgcag		
<pre></pre>	cagccaggco	ctgcag		36
<pre></pre>	cagccaggco	ctgcag		36
<pre></pre>	cagccaggco	ctgcag		36
<pre></pre>	cagecagged	ctgcag		36
<pre></pre>	cagecagged	ctgcag		36
<pre></pre>	cagecagged	ctgcag		36
<pre></pre>	cagccaggcc	c ctgcag		36 36
<pre></pre>	cagccaggcc	c ctgcag		36
<pre></pre>	cagccaggcc	c ctgcag		36 36
<pre></pre>	cagccaggcc	c ctgcag		36 36

<212> DNA <213> Synthetic		
<400> 67		
cgcgcgcaca tgtctcaggc cctgcagctg gcaggc		36
<210> 68		
<211> 36		
<212> DNA		
<213> Synthetic		
. <b>&lt;4</b> 00> 68		7.5
egegegeaca tgtetgeect geagetggea ggetge		3.6
<210> 69	•	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 69		
cgcgcgcaca tgtctctgca gctggcaggc tgcttg		36
<210> 70		
<211> 36		
<212> DNA		
<213> Synthetic		٠.
<400> 70	•	
cgcgcgcaca tgtctcagct ggcaggctgc ttgagc		. 36
<210> 71		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 71		36
cgcgcgcaca tgtctctggc aggctgcttg agccaa		30
<210> 72		
<211> 36	. •	
<212> DNA		
<213> Synthetic		
<400> 72	•	
cgcgcgcaca tgtctgcagg ctgcttgagc caactc	·	36
<210> 73		
<211> 36		
<212> DNA	•	
<213> Synthetic		
<4G2> 73		
cacacaca rareradera errasderas ercesr		36
<210> 74		
<211> 36		
<212> DNA		
<213> Synthetic		
	•	
<400> 74 .		36
cgcgcgcaca tgtcttgctt gagccaactc catagc		36
<210> 75		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 75		
cgcgcgcaca tgtctttgag ccaactccat agcggc		36

<210> 76 <211> 36				-	
<212> DNA <213> Synthetic					
<400> 76 cgcgcgcaca tgtctagcca	actccatagc	ggcctt			36
<210> 77 <211> 36					
<212> DNA <213> Synthetic					
<400> 77 cgcgcgcaca tgtctcaact	ccatagcggc	ctttc			36
<210> 78 <211> 36 <212> DNA					
<213> Synthetic <400> 78					
cgcgcgcaca tgtctctcca	tagcggcctt	tteete	•		36
<210> 79 <211> 36 <212> DNA				•	
<213> Synthetic <400> 79			٠		
cgcgcgcaca tgtctcatag	cggccttttc	ctctac	-		.36
<211> 36 <212> DNA <213> Synthetic	•			·	
<400> 80 cgcgcgcaca tgtctagcgg	cettttcctc	taccag	•	•	36
<210> 81 <211> 36 <212> DNA <213> Synthetic					
<400> 81 cgcgcgcaca tgtctggcct	tttcctctac	: cagggg			36
<210> 82 <211> 36 <212> DNA <2:3> Synthetic					
<400> 82 c <i>gcgcg</i> ceca tgtctcttt	cctctaccaç	gggctc			36
<210> 83 <211> 36 <212> DNA <213> Synthetic					
<400> 83 cgcgcgcaca tgtctttcct	ctaccaggg	ctcctg			36
<210> 84 <211> 36 <212> DNA <213> Synthetic	·				

<400> 84		. •	
cgcgcgcaca tgtctctcta cca	aggggctc ctgcag		36
<210> 85			
<211> 36			
<212> DNA			
<213> Synthetic			
<400> 85			
cgcgcgcaca tgtcttacca gg	ggctcctg caggcc		36
<210> 86			
<211> 36			
<212> DNA			
<213> Synthetic			
<400> 86			
cgcgcgcaca tgtctcaggg gc	tectgeag geeetg		36
cgcgcgcaca cgccccaggg s			
<210> 87			
<211> 36			
<212> DNA			•
<213> Synthetic	•	•	
1,220 0,0000000			
<400> 87		•	
egegegeaca tgtetggget co	tgcaggcc ctggaa		3.6
<210> 88			
<211> 36	·		
<212> DNA			
<213> Synthetic			
<400> 88			
egegegeaca tgtetetect ge	aggccctg gaaggg		36
			•
<210> 89			
<211> 36			
<212> DNA			
<213> Synthetic		. •	
<400> 89			
cgcgcgcaca tgtctctgca gg	ecctggaa gggata		36
<210> 90			
<211> 36	•		
<212> DNA			
<213> Synthetic			
<400> 90			26
egegegeaca tgteteagge e	ctggaaggg atatco	:	36
<210> 9?			
<211> 36			
<212> DNA		÷	
<213> Synthetic		•	
		i	
<400> 91			36
cgcgcgcaca tgtctgccct g	gaagggata tcccc	•	30
<210> 92	•		
<211> 36			
<212> DNA			
<213> Synthetic			
<400> 92			26
cgcgcgcaca tgtctctgga a	gggatatecccsgaq	. ·	36

<210> 95

.211 - 26		
<211> 36	-	
<212> DNA		
<213> Synthetic		
<400> 93		
cgcgcgcaca tgtctgaagg gatatccccc gagttg		36
egegegeaca egeetgaagg gatatecooo gagtog		
·		
<210> 94		
<211> 36		
<212> DNA		
<213> Synthetic		
(213) Synchicoae	•	
<400> 94		20
cgcgcgcaca tgtctgggat atcccccgag ttgggt		36
<210> 95		
<211> 36		
<212> DNA		
<213> Synthetic		
	•	
<400> 95	•	75
cgcgcgcaca tgtctatatc ccccgagttg ggtccc		36
<210> 96		
<211> 36	•	
<212> DNA		
<213> Synthetic		
<400> 96		
cgcgcgcaca tgtcttcccc cgagttgggt cccacc		36
<210> 97		
<211> 36		
<212> DNA		
<213> Synthetic		
.400- 07		
<400> 97		36
cgcgcgcaca tgtctcccga gttgggtccc accttg	•	50
<210> 98		
<211> 36		
<212> DNA		
<213> Synthetic		
72137 Synthetic		
-400× 09		
<400> 98		36
cgcgcgcaca tgtctgagtt gggtcccacc ttggac		-
<210> 99		
<211> 36		
<212> DNA		
<213> Synthetic		
-100× 00		
<400> 99	•••	36
cgcgcgcaca tgtctttggg tcccaccttg gacaca	• •	٠,
<210> 100		
<211> 36	•	
<212> DNA		
<213> Synthetic		
122. Oj., 01.0		
<400> 100		
		36
egegegezea igietggtee eacettggae acaetg		٠,٠
<210> 101		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 101		
74000 242		

egegegeaca tgteteceae ettggacaca etgeag	-	36
	•	
<210> 102		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 102		
cgcgcgcaca tgtctacctt ggacacactg cagctg		36
<210> 103		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 103		
cgcgcgcaca tgtctttgga cacactgcag ctggac		36
cgcgcgcaca egueesss		
<210> 104		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 104	•	
cgcgcgcaca tgtctgacac actgcagctg gacgtc		36
<210> 105		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 105		36
cgcgcgcaca tgtctacact gcagctggac gtcgcc	*	36
<210> 106		
<211> 36		
<212> DNA		
<213> Synthetic	· ·	
<400> 106		
egegegeaca tgtetetgea getggaegte geegae		3.6
<210> 107 .	· .	
<211> 36	·	
<212> DNA	•	
<213> Synthetic	·	
<400> 107		
cgcgcgcaca tgtctcagct ggacgtcgcc gacttt		36
cgcgcgcacc cgcccccgc		
<210> 108		
<211> 36		
<212> DNA	•	
<213> Synthetic	•	
<400> 108		
cgcgcgcaca tgtctctgga cgtcgccgac tttgcc	•	36
<210> 109		
<211> 36		
<212> DNA		
<213> Synthetic	•	
<400> 109	·	
cgcgcgcaca totctgacgt cgccgacttt gccacc	•	36
<210> 113		
<211> 36		

<213> Synthetic		
(213) Synchecic	-	
<400> 110		
cgcgcgcaca tgtctgtcgc cgactttgcc accacc		36
<210> 111		
<211> 36	•	
<212> DNA		
<213> Synthetic		
<400> 111		
cgcgcgcaca tgtctgccga ctttgccacc accatc		36
<210> 112	•	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 112		
cgcgcgcaca tgtctgactt tgccaccacc atctgg		36
cgcgcgcaca cgcccgcccc cgccacacacacacacac		
<210> 113		
<211> 36		
<212> DNA		
<213> Synthetic	r	
<400> 113		
cgcgcgcaca tgtcttttgc caccaccatc tggcag		36
cycycycaca tyroracay careacae co		
<210> 114	•	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 114		
cgcgcgcaca tgtctgccac caccatctgg cagcag		36
ogogogoada ogotogoada aaaaaaaaa		
<210> 115		
<211> 36		
<212> DNA	. *	
<213> Synthetic		
<400> 115		
cgcgcgcaca tgtctaccac catctggcag cagatg		36
<210> 116		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 116		
cgcgcgcaca tgtctaccat ctggcagcag atggaa	•	36
	•	
-210- 117		
<210> 117		
<211> 36		
<211> 36 <212> DNA		
<211> 36		
<211> 36 <212> DNA <213> Synthetic	1	
<211> 36 <212> DNA <213> Synthetic <400> 117		36
<211> 36 <212> DNA <213> Synthetic		36
<211> 36 <212> DNA <213> Synthetic <400> 117		36
<211> 36		36
<211> 36 <212> DNA <213> Synthetic <400> 117 cgcgcgcaca tgtctatctg gcagcagatg gaagaa <210> 118		36
<211> 36 <212> DNA <213> Synthetic <400> 117 cgcgcgcaca tgtctatctg gcagcagatg gaagaa <210> 118 <211> 36		36
<211> 36 <212> DNA <213> Synthetic  <400> 117 cgcgcgcaca tgtctatctg gcagcagatg gaagaa  <210> 118 <211> 36 <212> DNA <213> Synthetic		36
<211> 36 <212> DNA <213> Synthetic  <400> 117 cgcgcgcaca tgtctatctg gcagcagatg gaagaa  <210> 118 <211> 36 <212> DNA		36

<210> 119						
<211> 36				-		
<212> DNA						
<213> Synthetic						
CZIS/ Synthetic						
<400> 119						
cgcgcgcaca tgtctcagca	gatggaagaa	ctggga				36
egegegeded egessedges	555					
<210> 120						
<211> 36						
<212> DNA						
<213> Synthetic	<u>.</u>					
\2137 by						
<400> 120						
cgcgcgcaca tgtctcagat	ggaagaactg	ggaatg				36
cgcgogocon carries						
<210> 121						
<211> 36						
<212> DNA						
<213> Synthetic	2					
		• •				
<400> 121						
cgcgcgcaca tgtctatgg	a agaactggga	atggcc	.00			36
<210> 122	÷			•		
<211> 36						
<212> DNA						
<213> Synthetic	c .					
<400> 122						
cgcgcgcaca tgtctgaag	a actgggaatg	geceet				3.6
	•					
<210> 123						
<211> 36						
<212> DNA						
<213> Syntheti	c					
<400> 123						
cgcgcgcaca tgtctgaac	t gggaatggc	c cctgcc .				. 36
			•			
<210> 124						
<211> 36	•					
<212> DNA		•				
<213> Syntheti	.c					
<400> 124						2.0
cgcgcgcaca tgtctctgg	g aatggcccc	t gccctg				36
		•				
<210> 125			•			
<211> 36						
<212> DNA						
<213> Syntheti	ic					
				4 .		
<400> 125						36
cgcgcgcaca tgtctggaa	at ggcccctgc	c ctgcag				30
<210> 126			i			
<211> 36						
<212> DNA	•					
<213> Synthet:	ic	•				
			•			
<400> 126					•	36
cgcgcgcaca tgtctatg	gc ccctgccct	g cagece				30
					•	
<210> 127			•			
<211> 36						
<212> DNA		•				
<213> Synthet	jc	•				

36
36
36
36
36
36
36
36
36
36
סכ
36
36
30
36
36
36
36
36
36
36
36

<212> DNA	
<213> Synthetic	
,	
<400> 136	
cgcgcgcaca tgtctgccat gccggccttc gcctct	36
CGCGcGcaca raceraceae accasaceae accasaceae accasaceae	
<210> 137	
<211> 36	
<212> DNA	
<213> Synthetic	·
400 100	
<400> 137	36
cgcgcgcaca tgtctatgcc ggccttcgcc tctgct	
<210> 138	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 138	36
cgcgcgcaca tgtctccggc cttcgcctct gctttc	
<210> 139	
<211> 36	
<212> DNA	
<213> Synthetic	
•	
<400> 139	36
cgcgcgcaca tgtctgcctt cgcctctgct ttccag	
<210> 140	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 140	36
cgcgcgcaca tgtctttcgc ctctgctttc cagcgc	•
040-141	
<210> 141	
<211> 36	
<212> DNA	
<213> Synthetic	
,	
<400> 141	36
cgegegeaca tgtctgeete tgettteeag egeegg	
<210> 142	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 142	3'6
cgcgcgcaca tgtcttctgc tttccagcgc cgggca	
<210> 143	
<211> 36	
<212> DNA	
<213> Synthetic	
.400. 142	
<400> 143	36
cgcgcgcaca tgtctgcttt ccagcgccgg gcagga	
<210> 144	
<211> 36	•
<212> DNA	
<213> Synthetic	
<400> 144	36
cgcgcgcaca tgtctttcca gcgccgggca ggaggg	- /

<210> 145				-	
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 145					
	ccaaaceaaa	aggar c			36
cgcgcgcaca tgtctcagcg	ccgggcagga	999900			
<210> 146					
<211> 36					
<212> DNA					
<213> Synthetic					
-400- 146					
<400> 146	~~~~~~~	atceta			36
cgcgcgcaca tgtctcgccg	990099999	geeeeg			
<210> 147					
<211> 36					
<212> DNA					
	•				
<213> Synthetic					
<400> 147					
	200200000	ataart			36
cgcgcgcaca tgtctcgggc	aggaggggcc	ctggct			30
<210> 148					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 148					
cgcgcgcaca tgtctgcagg	POUCOT CCT O	attact			36
cgcgcgcaca cgcccgcagg	0999910009	guagua			
<210> 149					
<211> 36					
<212> DNA					
<213> Synthetic					
75.5					
<400> 149					
cgcgcgcaca tgtctggagg	aatcctaatt	gctagc			36
		•			
<210> 150					
<211> 36					
<212> DNA					
<213> Synthetic				•	
	•				
<400> 150					
cgcgcgcaca tgtctggggt	cctaattact	agccat			36
<210> 151					
<211> 36					
<212> DNA					
<213> Synthetic	:				
- ,		•		•	
<400> 151					
cgcgcgcaca tgtctgtcct	ggttgctagc	catctg		•	36
	-		ı	•	
<210> 152			,		
<211> 36					
<212> DNA					
<213> Synthetic	:				
,					
<400> 152					
cgcgcgcaca tgtctctggt	tgctagccat	ctgcag			36
<210> 153					
<211> 36					
<212> DNA			•		
<213> Synthetic	:				

<400> 153 cgcgcgcaca tgtctgttgc	tagccatctg	cagagc	-		36
<210> 154 <211> 36					
<212> DNA <213> Synthetic					
<400> 154 cgcgcgcaca tgtctgctag	ccatctgcag	agcttc			36
<210> 155					
<211> 36				•	
<212> DNA					
<213> Synthetic					
<400> 155					
cgcgcgcaca tgtctagcca	tctgcagagc	ttcctg			36
<210> 156					
<211> 36					
<212> DNA <213> Synthetic					
				•	
<400> 156 cgcgcgcaca tgtctcatct	gcagagette	ctagag		•	3.6
	gcagagecee	ccggag			
<210> 157 <211> 36					
<212> DNA					
<213> Synthetic					
<400> 157					
cgcgcgcaca tgtctctgca	gagetteetg	gaggtg			3.6
<210> 158					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 158					26
cgcgcgcaca tgtctcagag	cttcctggag	grgreg			36
<210> 159					
<211> 36					
<212> DNA <213> Synthetic					
<400> 159 cgcgcgcaca tgtctagctt	cctaaraata	teatec			36
cycycaca cycctayccc	cccggaggc	, ccycac			
<210> 160					
<211> 36			•		
<212> DNA <213> Synthetic				e eg	
			i		
<400> 160 cgcgcgcaca tgtctttcct	ggaggtgte	taccoc			36
		, caccyc			
<210> 161 <211> 36					
<211> 36 <212> DNA					
<213> Synthetic					
<400> 161 cgcgcgcaca tgtctctgga	ggtgtcgts				36
	325366360	- cycytt			-
<210> 162			•		

<211> 36						•
<212> DNA						
				-		
<213> Synthetic						
<400> 162						
cgcgcgcaca tgtctgaggt	ntcataceae	gttcta				36
cacacaca caccadaa	9009000090	90000				•
<210> 163						
<211> 36				•		
<212> DNA	•					
·						
<213> Synthetic						
<400> 163						
egegegeaca tgtctgtgtc	gtaccgcgtt	ctacgc				36
				•		
.0.0						
<210> 164						
<211> 36						
<212> DNA						
<213> Synthetic						
vers synthetic						
<400> 164						
cgcgcgcaca tgtcttcgta	ccgcgttcta	cgccac				36
<210> 165						
<211> 36						
<212> DNA						
<213> Synthetic						
,						
.400- 155						
<400> 165						
cgcgcgcaca tgtcttaccg	cgttctacgc	cacctt				36
<210> 166						
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 166						
	tetaceceae	cttaca				36
cgcgcgcaca tgtctcgcgt	cctacgccac	ccegeg				
		•				•
<210> 167				٠.		
<211> 36	*			•		
<212> DNA						
<213> Synthetic						
(21)> Synthetic						
<400> 167						
cgcgcgcaca tgtctgttct	acgccacctt	gcgcag				36
		311				
-210- 160						
<210> 168						
<211> 36						
<212> DNA					•	
<213> Synthetic						
•						
<400> 168						
cgcgcgcaca tgtctctacg	ccaccttgcg	cagccc				3.6
•						
<210> 169						
<211> 36						
			į			
<212> DNA						
<213> Synthetic						,
<400> 169						
						3.6
cgcgcgcaca tgtctcgcca	CCTTGCCC	000000				3.0
	ccttgcgcag	cccgac	•			
	ccttgcgcag	cccgac	,			
<210> 170	ccttgcgcag	cccgac				-
<210> 170 <211> 36	ccttgcgcag	cccgac	·			
<211> 36	ccttgcgcag	cccgac	· .			
<211> 36 <212> DNA	,	cccgac	· ·		<i>:</i>	
<211> 36	,	cccgac			<i>:</i>	
<211> 36 <212> DNA	,	cccgac			<i>.</i>	

cgcgcgcaca tgtctcacct tgcgcagccc gacatg	-	36
<210> 171		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 171		36
cgcgcgcaca tgtctcttgc gcagcccgac atggct		30
<210> 172		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 172		36
egegegeaca tgtetgegea geeegacatg getaca		30
<210> 173	•	
<211> 36		
<212> DNA		
<213> Synthetic	·	
<400> 173		36
egegegeaca tgteteagee egacatgget acacca		30
<210> 174		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 174		26
cgcgcgcaca tgtctcccga catggctaca ccattg		36
<210> 175		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 175		
cgcgcgcaca tgtctcgcca ccttgcgcag cccact		36
<210> 176		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 176	*	
cgcgcgcaca tgtctcacct tgcgcagccc actagt		36
<210> 177		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 177		26
cgcgcgcaca tgtctcttgc gcagcccact agtcat	1	36
<210> 178	•	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 178		• •
cgcgcgcaca tgtctgcgca gcccactagt catcca		36
<210> 173		•
<211> 36		
<212> DNA		

<213> Synthetic				-		
<400> 179						
cgcgcgcaca tgtctcagcc	cactagtcat	ccacct				36
egegegeaca egeeceagee	caccagaca					
<210> 180						
<211> 36						
<212> DNA						
<213> Synthetic						
(213) Dynamotic						
<400> 180						
cgcgcgcaca tgtctcccac	tagtcatcca	cctatg				36
cycycycaea cyforeiai		_				
<210> 181						
<211> 36						
<212> DNA						
<213> Synthetic						
-						
<400> 181						
cgcgcgcaca tgtctcgcca	ccttgcgcag	cccggc				36
-						
<210> 182					•	
<211> 36				•		
<212> DNA						
<213> Synthetic	•					
<400> 182						36
cgcgcgcaca tgtctcacct	tgcgcagccc	ggcggc				30
<210> 183						
<211> 36	•		*			
<212> DNA						
<213> Synthetic						
-400× 103						
<400> 183	acsucceuu.	accaac				36
<400> 183 cgcgcgcaca tgtctcttgc	gcagcccggc	ggegge				36
cgcgcgcaca tgtctcttgc	gcagcccggc	ggegge				36 .
cgcgcgcaca tgtctcttgc	gcagcccggc	ggcggc				36 .
<pre>cgcgcgcaca tgtctcttgc &lt;210&gt; 184 &lt;211&gt; 36</pre>	gcagcccggc	ggegge				36 .
<pre>cgcgcgcaca tgtctcttgc &lt;210&gt; 184 &lt;211&gt; 36 &lt;212&gt; DNA</pre>		ggcggc				36 .
<pre>cgcgcgcaca tgtctcttgc &lt;210&gt; 184 &lt;211&gt; 36</pre>		ggcggc		·		36 .
<pre>cgcgcgcaca tgtctcttgc &lt;210&gt; 184 &lt;211&gt; 36 &lt;212&gt; DNA</pre>		ggegge				
<pre><gcgcgcaca <210="" tgtctcttgc=""> 184 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 184</gcgcgcaca></pre>	:					36
<pre><gcgcgcaca <210="" tgtctcttgc=""> 184 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</gcgcgcaca></pre>	:					
<pre><gcgcgcaca <210="" tgtctcttgc=""> 184 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 184</gcgcgcaca></pre>	:					
<pre><gcgcgcaca <210="" tgtctcttgc=""> 184 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 184 cgcgcgcaca tgtctgcgca</gcgcgcaca></pre>	:					
<pre>&lt;210&gt; 184 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 184 cgcgcgcaca tgtctgcgca &lt;210&gt; 185 &lt;211&gt; 36 &lt;212&gt; DNA</pre>	geeeggegge					
<pre>&lt;210&gt; 184 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 184 cgcgcgcaca tgtctgcgca &lt;210&gt; 185 &lt;211&gt; 36</pre>	geeeggegge					
<pre><gcgcgcaca <210="" tgtctcttgc=""> 184 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 184 cgcgcgcaca tgtctgcgca &lt;210&gt; 185 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</gcgcgcaca></pre>	geeeggegge					
<pre>cgcgcgcaca tgtctcttgc</pre>	. geeeggegge	ggetet				36
<pre><gcgcgcaca <210="" tgtctcttgc=""> 184 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 184 cgcgcgcaca tgtctgcgca &lt;210&gt; 185 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</gcgcgcaca></pre>	. geeeggegge	ggetet				
<pre></pre>	. geeeggegge	ggetet				36
<pre>cgcgcgcaca tgtctcttgc</pre>	. geeeggegge	ggetet				36
<pre></pre>	. geeeggegge	ggetet				36
<pre></pre>	c cggcggcggc	ggetet				36
<pre></pre>	c cggcggcggc	ggetet				36
<pre></pre>	c cggcggcggc	ggetet				36
<pre></pre>	c cggcggcggc	e ggetet	, , ,			36
<pre></pre>	c cggcggcggc	e ggetet				36
<pre></pre>	c cggcggcggc	e ggetet				36
<pre></pre>	c cggcggcggc	e ggetet				36
<pre></pre>	c cggcggcggc	e ggetet				36
<pre></pre>	c caacaacre	e ggetet				36
<pre></pre>	c caacaacre	e ggetet				36
<pre></pre>	c caacaacre	e ggetet				36
<pre>cgcgcgcaca tgtctcttgc</pre>	cadeadacre	ggetet ctgac	, and the second			36

<210> 188				
<211> 36			-	
<212> DNA				
<213> Synthetic				
•				
<400> 188				36
tatatatgcg gccgcagcca	tgtcacgcgt acgatt			
<210> 189				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 189				36
tatatatgcg gccgcagcca	tgtcagagcc gccgcc			50
<210> 190				
<211> 36				
<212> DNA				
<213> Synthetic				
	•		•	
<400> 190	•		•	36
tatatatgcg gccgctgtag	ccatgtcggg ctgcgc			36
<210> 191				
<211> 36				
<212> DNA	•			
<213> Synthetic	:			
	·			
<400> 191				26
tatatatgcg gccgctgtag	ccatgicacg cgtacg			36
<210> 192				
<211> 36				
<212> DNA				
<213> Synthetic	•			
7,22, 5,1,0,0				
<400> 192	•			
tatatatgcg gccgctgtag	g ccatgtcaga gccgcc			36
<210> 193				
<211> 36				
<212> DNA	•			
<213> Synthetic	C	* .	•	
<400> 193		,		
tatatatgcg gccgctggt	g tagecatgic gggcts	I		3.6
tatataty you good or you	,,	•		
<210> 194				
<211> 36				
<212> DNA				
<213> Synthetic	c			
<400> 194				
tatatatgcg gccgctggt	a teoccatoro acoco:			36
tatatatgeg geegetgge	g tageoniget meres.	•		
<210> 195		,	•	
<211> 36		,		
<212> DNA				
<213> Syntheti	c			
<400> 195		_		36
tatatatgcg gccgctggt	e rapecatore agage	<u>-</u>		
<210> 196	•			
<211> 36				
<212> DNA	•			
<213> Syntheti	.c			

<400> 196		
tatatatgcg gccgccaatg gcgtagccat gccggg	•	36
tatatatycy googetaaty yeyongotat yeess	•	
-2105 107		
<210> 197		
<211> 36		
<212> DNA		
<213> Synthetic		
•		
<400> 197		
tatatatgcg gccgccaatg gtgtagccat gtcacg		36
<210> 198		
<211> 36		
<211> DNA		
<213> Synthetic	* ·	
**		
<400> 198		36
tatatatgcg gccgccaatg gtgtagccat gtcaga		36
<210> 199		
<211> 36		
<212> DNA		
<213> Synthetic	·	
<400> 199		
	•	36
tatatatgcg gccgcgccca atggtgtagc catgtc		
0.0		
<210> 200		
<211> 36		
<212> DNA		
<213> Synthetic		
	·	
<400> 200		
tatatatgcg gccgcagggc ccaatggtgt agccat		36
· · · · · · · · · · · · · · · · · · ·		
<210> 201		
<211> 36		
<212> DNA		
<213> Synthetic		
7225 Dynamou-1	*	
<400> 201	. •	
		36
tatatatgcg gccgcggcag ggcccaatgg tgtagc		
<210> 202		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 202		
tatatatgcg gccgcgctgg cagggcccaa tggtgt	· ·	36
<210> 203		
<211> 36		
<212> DNA		
<213> Synthetic		
4.50	•	
<430> 203	, **	20
tatatatgeg geegeggage tggeagggee caatgg		36
<210> 204		
<211> 35		-
<212> DNA		
<213> Synthetic	· ·	
<400> 204		
		36
tatatatded deedeevadda vaeeddaad deecaa		20
.910. 907		
<210> 205		
<211> 36		

minnonin aun minnes i

-2125 DNA				
<212> DNA				
<213> Synthetic	•	-		
<400> 205				
atatatgcg gccgcggca gggagctggc agggcc			3	36
atatatgeg geegegggea gggageegge aggge				
<210> 206				
<211> 36				
<212> DNA				
<213> Synthetic	•			
<400> 206				
				3.6
tatatatgeg geegeetggg geagggaget ggeagg				
<210> 207				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 207				
tatatatgcg gccgcgctct ggggcaggga gctggc				36
tatatatgeg geegeeee ggggeeggg				
<210> 208				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 208				
tatatatgcg gccgcgaagc tctggggcag ggagct				36
Cacacacaca Bacacama				
000				
<210> 209				
<211> 36				
<212> DNA				
<213> Synthetic			•	
(213) Syncholic				
<400> 209				36
tatatatgcg gccgccagga agctctgggg caggga				30
<210> 210				•
•				
<211> 36				
<212> DNA				-
<213> Synthetic				
•				
<400> 210				
				36
tatatatgcg gccgcgagca ggaagctctg gggcag				
· ·				
<210> 211				
<211> 36				
<212> DNA	•			
<213> Synthetic				
<400> 211				
tatatatgcg gccgccttga gcaggaagct ctgggg				36
Catatatgeg geegeetiga geaggaaget eeggag				
<210> 212				
<211> 36				
<212> DNA		•		
	F 1.			
<213> Synthetic				
<400> 212				
tatatatgcg gccgcagact tgagcaggaa gctctg				36
"wronded dechorage canadadi in second		_		
		•		
<210> 213				
<211> 36				
<212> DNA				
		•		
<213> Synthetic				
		•		
<400> 213				_
tatatatgcg gccgctaaag acttgagcag gaagct				36

<210> 214		
<211> 36		
<212> DNA		
<213> Synthetic		
•		
<400> 214		
tatatatgcg gccgcctcta aagacttgag caggaa		36
,		
<210> 215		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 215		3.6
tatatatgcg gccgcttgct ctaaagactt gagcag	•	30
<210> 216		
<211> 36		
<212> DNA		
<213> Synthetic		
-400- 216	•	
<pre>&lt;400&gt; 216 tatatatgcg gccgccactt gctctaaaga cttgag</pre>	:	3.6
tatatatgeg geogetatet getetabaga tergas		
<210> 217		
<211> 36		٠.
<212> DNA		
<213> Synthetic		
<400> 217		
tatatatgeg geegeeetea ettgetetaa agaett		36
<210> 218		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 218		
tatatatgcg gccgccttcc tcacttgctc taaaga		36
<210> 219	•	
<211> 36		
<212> DNA	•	
<213> Synthetic		
•		
<400> 219		36
tatatatgcg gccgcgatct tectcacttg ctctaa		30
<210> 220	•	
<210> 220 <211> 36		
<212> DNA		
<213> Synthetic		
<400> 220		
tatatatgcg gccgcctgga tcttcctcac ttgctc		36
	,	
<310> 221	·	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 221		37
tatatatgcg gccgcgccct ggatcttcct cacttg		36
	•	
<210> 222		
<211> 36		
	•	
<212> DNA <213> Synthetic		

מאוסחרותי אווח היופטובאיו

<400> 222	· -	
tatatatgcg gccgcatcgc cctggatctt cctcac		36
tatatated accaerage constraint		
<b>&lt;210&gt; 223</b>		
<211> 36		
<212> DNA		
<213> Synthetic		
(213) Synchicold		
<400> 223		
tatatatgcg gccgcgccat cgccctggat cttcct		36
<210> 224		
	•	
<211> 36		
<212> DNA		
<213> Synthetic	· ·	
<400> 224		
		36
tatatatgcg gccgctgcgc catcgccctg gatctt		30
•	•	
<210> 225		
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 225		
tatatatgcg gccgccgctg cgccatcgcc ctggat		36
Catalacycy geogeogeog ogolatogot coggan		
<210> 226		
<211> 36		
<212> DNA		
<213> Synthetic	•	
\2137 by\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\		
<400> 226		36
tatatatgcg gccgcgagcg ctgcgccatc gccctg		36
		•
<210> 227		
<210> 227 <211> 36		•
<210> 227 <211> 36 <212> DNA		•
<210> 227 <211> 36		•
<210> 227 <211> 36 <212> DNA		•
<210> 227 <211> 36 <212> DNA <213> Synthetic		•
<210> 227 <211> 36 <212> DNA <213> Synthetic <400> 227		36
<210> 227 <211> 36 <212> DNA <213> Synthetic		36
<210> 227  <211> 36  <212> DNA  <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc		36
<210> 227 <211> 36 <212> DNA <213> Synthetic <400> 227		36
<210> 227  <211> 36  <212> DNA  <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc		36
<210> 227		36
<210> 227		36
<210> 227		36
<210> 227		36
<210> 227		
<210> 227		36
<210> 227		
<210> 227     <211> 36     <212> DNA     <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc  <210> 228     <211> 36     <212> DNA     <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc		
<210> 227		
<210> 227		
<pre>&lt;210&gt; 227</pre>		
<210> 227		
<pre>&lt;210&gt; 227</pre>		
<pre>&lt;210&gt; 227</pre>		
<210> 227		
<pre>&lt;210&gt; 227</pre>		36
<210> 227     <211> 36     <212> DNA     <213> Synthetic      <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc      <210> 228     <211> 36     <212> DNA     <213> Synthetic      <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc      <210> 228     <211> 36     <212> DNA     <213> Synthetic      <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc      <210> 229     <211> 36     <212> DNA     <213> Synthetic      <400> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc		36
<210> 227		36
<210> 227     <211> 36     <212> DNA     <213> Synthetic      <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc      <210> 228     <211> 36     <212> DNA     <213> Synthetic      <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc      <210> 228     <211> 36     <212> DNA     <213> Synthetic      <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc      <210> 229     <211> 36     <212> DNA     <213> Synthetic      <400> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc		36
<210> 227 <211> 36 <212> DNA <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc  <210> 228 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 228 c211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc  <210> 230		36
<210> 227 <211> 36 <212> DNA <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc  <210> 228 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc  <210> 230 <211> 36 <212> DNA		36
<210> 227 <211> 36 <212> DNA <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc  <210> 228 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc  <210> 230 <211> 36		36
<210> 227 <211> 36 <212> DNA <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc  <210> 228 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 229 tatatatgcg gccgcctcct cctggagcgctgc gccatc  <400> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc  <210> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc  <210> 230 <211> 36 <212> DNA <213> Synthetic		36
<210> 227 <211> 36 <212> DNA <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc  <210> 228 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc  <210> 230 <211> 36 <212> DNA		36
<210> 227 <211> 36 <212> DNA <213> Synthetic  <400> 227 tatatatgcg gccgcctgga gcgctgcgcc atcgcc  <210> 228 <211> 36 <212> DNA <213> Synthetic  <400> 228 tatatatgcg gccgcctcct ggagcgctgc gccatc  <210> 229 <211> 36 <212> DNA <213> Synthetic  <400> 229 tatatatgcg gccgcctcct cctggagcgctgc gccatc  <400> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc  <210> 229 tatatatgcg gccgccttct cctggagcgc tgcgcc  <210> 230 <211> 36 <212> DNA <213> Synthetic		36
<pre>&lt;210&gt; 227</pre>		36
<pre>&lt;210&gt; 227</pre>		36

<211> 36	_
<212> DNA	<u>-</u>
<213> Synthetic	
(225° 6)	
.400- 221	
<400> 231	-ac 36
tatatatgcg gccgcacaca gcttctcctg gage	igc 50
<210> 232	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 232	
tatatatgcg gccgcggcac acagettete etgg	gag 36
<210> 233	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 233	
tatatatgcg gccgcggtgg cacacagctt ctc	ctg 36
	,
<210> 234	
<del></del>	· · · · · · · · · · · · · · · · · · ·
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 234	•
tatatatgcg gccgcgtagg tggcacacag ctt	ctc 36
<210> 235	
•	•
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 235	••
tatatatgcg gccgccttgt aggtggcaca cag	ictt 36
	•
<210> 236	•
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 236	
tatatatgcg gccgccagct tgtaggtggc aca	acag 36
<210> 237	
<211> 36	• .
<212> DNA	
<213> Synthetic	
·	
<400> 237	
tatatatgcg gccgcgcaca gcttgtaggt gg	caca 36
•	
<210> 238	
<211> 36	1
<212> DNA	
<213> Synthetic	
<400> 238	
tatatatgeg geegegtgge acagettgta gg	tggc 36
<210> 239	
<211> 36	•
<212> DNA	
<213> Synthetic	

<400> 239

tatatatgcg gccgcgtgtc cgagcagcac cagctc

<210> 248
<211> 36
<212> DNA

36

<213> Synthetic				
<400> 248		•	•	
tatatatgcg gccgcagagt	gtccgagcag	caccag		36
<210> 249				
<211> 36			•	
<212> DNA				
<213> Synthetic				
<400> 249				
tatatatgcg gccgccagag	agtgtccgag	cagcac		36
		-		
<210> 250				
<211> 36 <212> DNA				
<213> Synthetic				
<400> 250	asasatatac	gaggag		3.6
tatatatgcg gccgcgccca	gagagegeee	gagcag		30
<210> 251				
<211> 36			•	
<212> DNA				
<213> Synthetic	•		•	
<400> 251			•	
tatatatgcg gccgcgatgc	ccagagagtg	tccgag		36
<210> 252				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 252				
tatatatgcg gccgcgggga	tgcccagaga	gtgtcc		36
<210> 253 <211> 36				
<212> DNA		•		
<213> Synthetic				
-400× 2E2				
<400> 253 tatatatgcg gccgcccagg	ggatgccag	agagtg		36
<210> 254				
<211> 36				
<212> DNA <213> Synthetic				
\213> Synthetic			·	
<400> 254				
tatatatgcg gccgcagccc	aggggatgcc	cagaga		36
<210> 255				
<211> 36				
<212> DNA				
<213> Synthetic			· · ·	
<400> 255				
tatatatgcg gccgcgggag	cccaggggat	gcccag		36
<210> 256 <211> 36				
<211> 36 <212> DNA				
<213> Synthetic				
<400> 256		ant	•	3.e
tatatatgcg gccgccaggg	yaycccaygg	Aaracc		36

<210> 257		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 257		
		36
tatatatgcg gccgcgctca ggggagccca ggggat		
010: 050		
<210> 258		
<211> 36 .		
<212> DNA		
<213> Synthetic		
	-30-	
<400> 258		36
tatatatgcg gccgcggagc tcaggggagc ccaggg		30
	•	
<210> 259		
<211> 36		
<212> DNA		
<213> Synthetic		
•		
<400> 259	•	
tatatatgcg gccgcgcagg agctcagggg agccca		36
cacacacycy googogoass as a second		
<210> 260		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 260		36
tatatatgcg gccgcggggc aggagctcag gggagc		
•	· ·	
<210> 261		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 261		
tatatatgcg gccgcgctgg ggcaggagct cagggg		36
<210> 262		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 262		
tatatatgcg gccgcctggc tggggcagga gctcag		36
tutututgeg geograpsga -33355 - 2		
<210> 263		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 263		36
tatatargeg geegeggeet ggerggggea ggaget		
<210> 264	ı	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 264		_
tatatatgcg gccgccaggg cctggctggg gcagga		36
<210> 265		
<211> 36		
<212> DNA		
<213> Synthetic		

3€

THISDOOID AND COLORES I

<400> 273

<210> 274 <211> 36

tatatatgcg gccgcgaggt tgctcaagca gcctgc

<212> DNA <213> Synthetic			-	-		
<400> 274 tatatatgcg gccgcatgga g	ggttgctcaa	gcagcc				36
<210> 275 <211> 36						
<211> 36 <212> DNA						
<213> Synthetic						
<400> 275						
tatatatgcg gccgcgctat	ggaggttgct	caagca				36
<210> 276						
<211> 36						
<212> DNA <213> Synthetic						
(213) Synthetic						
<400> 276	•					
tatatatgcg gccgcgccgc	tatggaggtt	gctcaa				36
					•	
<210> 277						
<211> 36	•					
<212> DNA <213> Synthetic						
(213) Synthetic						•
<400> 277						
tatatatgcg gccgcaaggc	cgctatggag	gttgct				36
<210> 278						
<211> 36				•		
<212> DNA <213> Synthetic	•					
(213) Bynchecic					. •	
<400> 278						
tatatatgcg gccgcgaaaa	ggccgctatg	gaggtt				36
<210> 279 <211> 36						
<211> 36 <212> DNA						
<213> Synthetic						
•	•					
<400> 279						36
tatatat <del>go</del> g googogagga	aaaggccgct	atggag				30
<210> 280						
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 280		cerata				36
tatatatgcg gccgcgtaga	ggaaaaggc	. yctaty				
<210> 281						
<211> 36						
<212> DNA			1			
<213> Synthetic	•		•			
-440- 201-		•				
<400> 281 tatacatgeg geogeotget	AGAGGAAAA	a accast				36
rarararach accederión.	-And Annual	33				
<210> 282						
<211> 36						
<212> DNA					•	
<213> Synthetic					2	
<400> 282						
tatatatgcg gccgcccct	ggtagagga	a aaggcc				36

<210> 283			-		
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 283					
tatatatgcg gccgcgagcc	cctggtagag	gaaaag			36
catatacycy geograps		J			
<210> 284					
<211> 36					
<212> DNA					
•					
<213> Synthetic					
<400> 284					
tatatatgcg gccgccagga	gcccctggta	gaggaa			36
-210- 205					
<210> 285					
<211> 36					
<212> DNA					
<213> Synthetic					
4225					
100 005					
<400> 285					
tatatatgcg gccgcctgca	ggagcccctg	gtagag			36
				•	
<210> 286					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 286					
tatatatgcg gccgcggcct	gcaggagece	ctoota			36
tatatatyty geografier	9003303000	005500			
<210> 287					
<211> 36					
<212> DNA					
<213> Synthetic					
(21)/ Synthetic					
<400> 287					
tatatatgcg gccgccaggg	cctgcaggag	cccctg	•		3.6
<210> 288					
<211> 36					
<212> DNA					
<213> Synthetic					
-					
<400> 288					
		~~~~~			36
tatatatgcg gccgcttcca	gggcctgcag	gageee			
<21 <b>0&gt;</b> 289			•		
<211> 36					
<212> DNA					
<213> Synthetic					
			•		
<400> 289					
tatatatgcg gccgcccctt	ccagggcctg	caggag			36
				•	
<210> 290			i		
<211> 36					
<212> DNA					
<213> Symthetic			•		
<400> 290					
					٠,٠
tatatatgcg gccgctatcc	cttccagggc	cracad			36
<210> 291					
<211> 36	• ,				
<212> DNA					
<213> Synthetic					

<400> 291 tatatatgcg gccgcggata tcccttccag ggcctg	36
<210> 292	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 292	36
tatatatgcg gccgcggggg atatcccttc cagggc	30
<210> 293	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 293	3.6
tatatatgcg gccgcctcgg gggatatccc ttccag	30
<210> 294	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 294	26
tatatatgcg gccgccaact cgggggatat cccttc	36
<210> 295	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 295	
tatatatgcg gccgcaccca actcggggga tatccc	3.6
<210> 296	
<211> 36	
<212> DNA	·
<213> Synthetic	
<400> 296	36
tatatatgcg gccgcgggac ccaactcggg ggatat	36
<210> 297	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 297	
tatatatgeg geegeggtgg gacceaacte ggggga	36
<210> 29&	
<211> 36	
<212> RVA	
<213> Synthetic	
<400> 298	•
tatatætgeg geegeeaagg tgggaeeeaa eteggg	36
<210> 299	
<211> 36	
<212> DNA	
<213> Symthetic	
<400> 299	
tatatatgeg geogegteea aggragaee caacte	36
<210> 300	

<211> 36 <212> DNA	-
<213> Synthetic	
22137 Synthetic	
<400> 300	
tatatatgeg geegetgtgt ccaaggtggg acccaa	3.6
<210> 301	
<211> 36	
<212> DNA	
<213> Synthetic	
.400. 201	•
<400> 301	36
tatatatgcg gccgccagtg tgtccaaggt gggacc	50
<210> 302	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 302	
tatatatgcg gccgcctgca gtgtgtccaa ggtggg	36
	$\cdot$
<210> 303	
<211> 36	•
<212> DNA	•
<213> Synthetic	•
<400> 303	36
tatatatgcg gccgccagct gcagtgtgtc caaggt	30
<210> 304	
<211> 36	
<211> 30 <212> DNA	
<213> Synthetic	
<400> 304	
tatatatgcg gccgcgtcca gctgcagtgt gtccaa	36
<210> 305	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 305	
tatatatgcg gccgcgacgt ccagctgcag tgtgtc	36
tatatatata accaesacat comandana alama	
<210> 306	
<211> 36	: .
<212> DNA	
<213> Synthetic	
<400> 306	
tatatatgcg gccgcggcga cgtccagctg cagtgt	. 36
210. 207	
<210> 307	
<211> 36 <212> DNA	i
<212> SNA <213> Synthetic	
-eas- ojmenedat	•
<480> 307	
ratatatgeg geogegicegg egacgiceag eigeag	36
<210> 308	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 308	

atatatgcg gccgcaaagt cggcgacgtc cagctg	36
<210> 309	
<211> 36	
<212> DNA	
<213> Synthetic	
. <400> 309	26
tatatatgcg gccgcggcaa agtcggcgac gtccag	36
.010- 310	
<210> 310 <211> 36	
<212> DNA	
<213> Synthetic <400> 310	
<400> 310	
tatatatgcg gccgcggtgg caaagtcggc gacgtc	36
<210> 311	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 311	
tatatatgcg gccgcggtgg tggcaaagtc ggcgac	36
tatatatgeg geogegguas asserbas assessment	
<210> 312	
<211> 36	
<212> DNA	
<213> Synthetic	
*	
<400> 312	36
tatatatgcg gccgcgatgg tggtggcaaa gtcggc	30
<210> 313	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 313	
tatatatgcg gccgcccaga tggtggtggc aaagtc	36
<210> 314	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 314	
tatatatgcg gccgcctgcc agatggtggt ggcaaa	36
<210> 315	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 315	
teracating decoderact decadarate adrage	36
es att case geographe geographe go ob ob o	
<210> 316	
<211> 36	
<212> DNA	
<pre>&lt;313&gt; Synthetic</pre>	
<400> 316	36
tatatatgcg gccgccatct gctgccagat ggtggt	
<210> 317	•
<211> 36	
<211 > 30	

<211> 36						
<212> DNA						
<213> Synthetic						•
<400> 231	•					
tatatatgcg gccgcacaca	gettereeta i	nancoc				36
tatatatgeg geograciaea	generates	909090				
-2105 222						
<210> 232						
<211> 36						
<212> DNA		·				
<213> Synthetic						
<400> 232			•			26
tatatatgcg gccgcggcac	acagettete	ctggag				36
	•					
<210> 233						
<211> 36			*			
<212> DNA						
<213> Synthetic	-					
	•					
<400> 233						
tatatatgcg gccgcggtgg	cacacagctt	ctcctg				36
<210> 234						
<211> 36	•				4	
<212> DNA			•			
<213> Synthetic	!					
<400> 234						
tatatatgcg gccgcgtagg	tggcacacag	cttctc				36
•						
<210> 235						
<211> 36						
<212> DNA	•					
<213> Synthetic	•					
<400> 235						
tatatatgcg gccgccttgt	aggtggcaca	cagctt				36
<210> 236			•			
<211> 36						
<212> DNA		•				
<213> Synthetic	•					
<400> 236						
tatatatgcg gccgccagct	tgtaggtggc	acacag				36
•						
<210> 237				•		
<211> 36						
<212> DNA						
<213> Synthetic	•					
•						
<400> 237						
tatatatgcg gccgcgcact	a gcttgtaggt	ggcaca	•			.36
			•			
<210> 238						
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 238			•			
tatatatgcg gccgcgtgg	c acagcttgta	ggtggc				3.6
<210> 239			•			
<211> 36						
<212> DNA						
<213> Synthetic	c ·				.•	
<400> 239	. •	•				

tatatatgcg geegegggt ggeacagett gtaggt		30
<210> 240	·	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 240		36
tatatatgcg gccgcctcgg ggtggcacag cttgta		
<210> 241		
<211> 36		
<212> DNA	A Committee of the Comm	
<213> Synthetic	•	
	•	
<400> 241	· .	36
tatatatgeg geegeeteet eggggtggea cagett	• .	
<210> 242		
<211> 36		
<212> DNA		
<213> Synthetic	• · · · · · · · · · · · · · · · · · · ·	
<400> 242 tatatatgcg gccgccagct cctcggggtg gcacag	•	36
tatatatgcg gccgccagcc ccccaaaaaaa		
<210> 243		
<211> 36	·	
<212> DNA	•	
<213> Synthetic		•
<400> 243		
tatatatgcg gccgccacca gctcctcggg gtggca	•	36
<210> 244		
<211> 36 <212> DNA		
<213> Synthetic		
· ·		
<400> 244	. •	36
tatatatgeg geegeeagea ceageteete ggggtg		,,,
<210> 245		
<211> 245 <211> 36		
<211> JONA		
<213> Synthetic	•	
•		
<400> 245	. •	36
tatatatgcg gccgcgagca gcaccagctc ctcggg	•	
<210> 246		
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 246		
tatatatgcg gccgctccga gcagcaccag ctcctc		36
66666669		
<210> 247		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 247		
tatatatgeg geogegigte egageageae cagete		36
	·	
<210> 248		
<211> 36 <212> DNA		
. <612- PM		

<213> Synthetic				
<400> 248				
tatatatgcg gccgcagagt gt	ccgagcag	caccag		36
<210> 249				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 249				
tatatatgcg gccgccagag ag	gtgtccgag	cagcac		36
<210> 250			•	
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 250		anacna		36
tatatatgcg gccgcgccca ga	agagtgttt	gageag		30
<210> 251			•	
<211> 36				
<212> DNA				
<213> Synthetic				
•				•
<400> 251				
tatatatgcg gccgcgatgc co	cagagagtg	tccgag		36
			I-	
<210> 252			-	
<211> 36				
<212> DNA				
<213> Synthetic			• 08	
<400> 252				
tatatatgcg gccgcgggga t	gcccagaga	gtgtcc		36
<210> 253				
<211> 36		,		•
<212> DNA				
<213> Synthetic	•			
<400> 253				36
tatatatgcg gccgcccagg g	gatgeceag	agagtg		30
<210> 254				
<211> 254				
<211> 30 <212> DNA				
<213> Synthetic				
			•	
<400> 254			•	
tatatatgcg gccgcagccc a	ggggatgcc	cagaga		36
<210> 255				•
<211> 36				
<212> DNA				
<213> Synthetic			•	
<400> 255				
tatatatgcg gccgcgggag c	ccaggggat	accesa		36
		50-009		•
<210> 256			•	
<211> 36				
<212> DNA				
<213> Synthetic				•
<400> 256				
tatatatgcg gccgccaggg g	Jagcccaqqq	gatgcc		36

<210> 257						
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 257						
tatatatgcg gccgcgctca	ggggagccca	ggggat				36
Catatachen googestee	3000-5	5555				
<210> 258	•					
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 258						
tatatatgcg gccgcggagc	tcagggagc	ccaggg				36
<210> 259						
<211> 36						
<212> DNA						
<213> Synthetic						
<213> Synthetic	•					
<400> 259						
tatatatgcg gccgcgcagg	agctcagggg	agccca				36
-210- 260						
<210> 260						
<211> 36	•					
<212> DNA						
<213> Synthetic						
1215 bynemetre						
		•				
<400> 260						
tatatatgcg gccgcggggc	aggagctcag	gggagc				36
<210> 261						
<211> 36						
<212> DNA						
<213> Synthetic		•				1.
.400: 061						
<400> 261						
tatatatgcg gccgcgctgg	ggcaggagct	cagggg				36
<210> 262						
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 262						
						36
tatatatgcg gccgcctggc	tggggcagga	gctcag				30
<210> 263						
<211> 36						
<212> DNA						
<213> Synthetic	:				•	
*						
<400> 263						
		aasaat				36
tatatatgcg gccgcggcct	. ggcrggggca	ggager				
<210> 264						
<211> 36						
<212> DNA						
<213> Synthetic	•					
<400> 264			•			
					•	36
tatatatgcg gccgccaggg	, cerageragg	, acayya				
<210> 265					٠.	
<211> 36						
∠212~ DAIA				•	•	
<212> DNA <213> Synthetic					·	

<400> 265 .	
tatatatgcg gccgcctgca gggcctggct ggggca	36
<210> 266	
<211> 36	•
<212> DNA	
<213> Synthetic	
(213) Synchistra	
1400- 355	
<400> 266	36
tatatatgcg gccgccagct gcagggcctg gctggg	
222	
<210> 267	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 267	36
tatatatgcg gccgctgcca gctgcagggc ctggct	. 30
<210> 268	
<211> 36	
<212> DNA	
<213> Synthetic	
•	
<400> 268	
tatatatgcg geogegeetg coagetgcag ggeetg	36
<210> 269	
<211> 36	
<2112 DNA	
<213> Synthetic	
<513> Symmetre	
<400> 269	
tatatatgcg gccgcgcagc ctgccagctg cagggc	36
tatatatgeg geogrape engulations only	
22.05 270	
<210> 270	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 270	36
tatatatgcg gccgccaagc agcctgccag ctgcag	
<210> 271	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 271	36
tatatatgog googogotoa agcagootgo cagotq	30
<210> 272	
<211> 36	
<212> DRA	
<213> Synthetic	
<400> 272	
tatatatgeg geogegitge teaageagee tgecag	36
<210> 273	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 273	
tatatatgcg gccgcgaggt tgctcaagca gcctgc	3 €
rarararach Acchendae caesamaen asses.	
-210> 274	
<210> 274	

<212> DNA						
<213> Synthetic			•			
<400> 274						
tatatatgcg gccgcatgga	aattactcaa	geagee				36
tatatacycy getycatyga	ggccgcccaa	900900	•			
-210. 275						
<210> 275						
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 275						26
tatatatgcg gccgcgctat	ggaggttgct	caagca				36
<210> 276						
<211> 36						
<212> DNA						
<213> Synthetic	!					
	•					
<400> 276		. •				
tatatatgcg gccgcgccgc	tatggaggtt	gctcaa				36
					•	
<210> 277				:		
<211> 36						
<212> DNA						
<213> Synthetic	:					
•••••						
<400> 277						
tatatatgcg gccgcaaggc	coctatoga	attact				36
<210> 278						
<211> 36						
<212> DNA						
<213> Synthetic	•					
(213) Synchectic						
<400> 278					•	
tatatatgcg gccgcgaaaa	gaccactato	gaggtt				36
catatatyty getychaat	, ggccgccac	. 505500	÷			
<210> 279						
<211> 36						
<212> DNA						
<213> Synthetic	-				•	
(213) Synchectic	•					
<400> 279	•			·		
		at conc				36
tatatatgeg geogegagga	a aaayyeeye	. acygag				-
-210- 250						
<210> 280				•		
<211> 36						
<212> DNA	_					
<213> Synthetic	5					
<400> 280						3.6
tatatatgcg gccgcgtag	a ggaaaaggc	gctatg				30
<210> 281						
<211> 36						
<212> DNA						
<213> Synthetic	C					
<400> 281				•		
tatacatgog geogooteg	agaggaaaa	geeget				36
· ·				,		
<210> 282			•			
<211> 36						•
<212> DNA						
<213> Syntheti	c				•	
<400> 282				_	•	
taratatoco occocccc	t ggtagagga	a aegge		••	·	36

<210> 283	
<211> 36	
<212> DNA	
<213> Synthetic	
· • • • • • • • • • • • • • • • • • • •	
<400> 283	
tatatatgcg gccgcgagcc cctggtagag gaaaag	3.6
Catatatgey geogegagee ceeggeagag gaaaag	
-210> 204	•
<210> 284	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 284	
tatatatgcg gccgccagga gcccctggta gaggaa	36
<210> 285	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 285	
tatatatgcg gcegcctgca ggagcccctg gtagag	36
<210> 286	•
<211> 36	•
<211> 50 <212> DNA	
<213> Synthetic	
<400> 286	36
tatatatgcg gccgcggcct gcaggagccc ctggta	36
***	
<210> 287	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 287	
tatatatgcg gccgccaggg cctgcaggag cccctg	36
<210> 288	
<211> 36	
<212> DNA	
<213> Synthetic	
•	
<400> 288	
tatatatgcg gccgcttcca gggcctgcag gagccc	36
<210> 289	
<211> 36	. 44
<212> DNA	
<213> Synthetic	
(21) Synthetic	
4005 200	
<400> 289	. 36
tatatatgeg geegeeeett eeagggeetg caggag	3.
-2105 200	
<210> 290	
<211> 36	
<212> DNA	
<213> Symthetic	
<400> 290	
tatatatgcg gccgctatcc cttccagggc ctgcag	36
<210> 291	
<211> 36	
<212> DNA	

<400> 291					
tatatatgcg gccgcggata	tecettecag	ggcctg			36
<210> 292					
<211> 36					
<212> DNA					
<213> Synthetic		,			
<400> 292					
tatatatgcg gccgcggggg	atatcccttc	cagggc			36
<210> 293					
<211> 36					
<212> DNA					
<213> Synthetic					
_					
<400> 293					
tatatatgcg gccgcctcgg	gggatatccc	ttccag			36
catacategog googeoco	333				
<210> 294					
<211> 36	•				
<212> DNA			,		
<213> Synthetic					
<400> 294			•		2-
tatatatgcg gccgccaact	cgggggatat	cccttc			36
<210> 295					
<211> 36			•		
<212> DNA					
<213> Synthetic					
<400> 295					
tatatatgcg gccgcaccca	actcggggga	tatece			36
<210> 296					
<211> 36					
<212> DNA					•
<213> Synthetic	!		•		
			•		
<400> 296					
tatatatgcg gccgcgggac	ccaactcgg	ggatat			36
taracathch Acchedagae		, ,,,			
-210- 207					
<210> 297			·		
<211> 36					
<212> DNA					
<213> Synthetic	•				
<400> 297					3.6
tatatatgcg gccgcggtgg	gacccaacto	: ggggga			36
<210> 298					
<211> 36					
<:112> DNA					
<213> Synthetic	:		•	•	
<400> 298					
tetatatgcg gccgccaagg	tgggacccaa	a ctcggg	•		36
<210> 299					
<211> 36			. •		
<212> DNA					
<21.3> Symthetic	•		•		
	-				
<400> 299					
tatatatgcg gccgcgtcca	. Addtascec	CARCCO			36
there ended Accardence					50

<211> 36	
<212> DNA	
<213> Synthetic	
<400> 300	
tatatatgcg gccgctgtgt ccaaggtggg acccaa	36
<210> 301	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 301	36
tatatatgcg gccgccagtg tgtccaaggt gggacc	30
<210> 302	
<211> 36	
<212> DNA	
<213> Synthetic	
•	
<400> 302	
tatatatgcg gccgcctgca gtgtgtccaa ggtggg	36
Lacacacaca accacacaca and accasacacacacacacacacacacacacacacacacac	
<210> 303	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 303	36
tatatatgcg gccgccagct gcagtgtgtc caaggt	70
<210> 304	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 304	
tatatatgcg gccgcgtcca gctgcagtgt gtccaa	. 36
<210> 305	
<211> 36	
<212> DNA	
<213> Synthetic	
· V213/ Synthetic	
4005 205	
<400> 305	3.6
tatatatgcg gccgcgacgt ccagctgcag tgtgtc	
<210> 306	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 306	
tatatatgcg gccgcggcga cgtccagctg cagtgt	3.6
<210> 307	
<211> 36	
<212> INA	
<213> Synthetic	
<480> 307	
	36
rataratgeg geogegiogg ogaogicoag otgoag	-
.010- 200	
<210> 308	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 308	

1	tatatatgeg geegeaaagt	cggcgacgtc	cagctg			36
	<210> 309					
	<211> 36					
	<212> DNA					
	<213> Synthetic					
	<400> 309					
	tatatatgcg gccgcggcaa	agtcggcgac	gtccag	•		36
	<210> 310					
	<211> 36					
	<212> DNA					
	<213> Synthetic					
	<400> 310					
	tatatatgcg gccgcggtgg	CABAGTEGGE	gacgtc			36
	Lacacacycy googogy					
	<210> 311					
	<211> 36					
	<212> DNA					
	<213> Synthetic					
	-			.*		
	<400> 311					
	tatatatgcg gccgcggtgg	tggcaaagtc	ggcgac	•		36
					8	
	<210> 312					
	<211> 36					
	<212> DNA					
	<213> Synthetic	ŀ		•		
	<400> 312					
	tatatatgcg gccgcgatgg	rogtogcaaa	atcaac			36
	catatatyty yeeguuyy	C3309944	555-			
	<210> 313					
	<211> 36			*		
	<212> DNA					
	<213> Synthetic	;				
		•				
	<400> 313			•		
	tatatatgcg gccgcccaga	tggtggtgg	aaagtc			36
	<210> 314					
	<211> 36					
	<212> DNA	_				
	<213> Synthetic	•				
	<400> 314					
	tatatatgcg gccgcctgcc	agatogtog	ggcaaa			36
	<210> 315					
	<211> 36					
	<212> DNA					
	<213> Synthetic	2		•		
	<400> 315					26
	tetacatgog googcotgo	t gccagatgg	r garaac			36
	-210- 316					
	<21C> 31€ <211> 36	•	•			
	<211> 36 <212> DNA					
	<pre>&lt;212&gt; bkk &lt;313&gt; Synthetic</pre>	2			•	
	-oad a mlosassage	=				
	<400> 316			•		
	tatatatgcg gccgccatc	t gctgccaga	ggtggt		•	36
					•	
	<210> 317	•				
	<211> 36					

<213> Synthetic					
<400> 317					
tatatatgcg gccgcttcca	tctgctgcca	gatggt			36
<210> 318					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 318					
tatatatgcg gccgcttctt	ccatctgctg	ccagat			36
<210> 319			•		
<211> 36		•		•	
<212> DNA					
<213> Synthetic					
<400> 319					
tatatatgcg gccgccagtt	cttccatctg	ctgcca			36
<210> 320			•		
<211> 36					
<212> DNA					
<213> Synthetic	•				
<400> 320		•			
tatatatgcg gccgctccca	gttettecat	ctgctg			36
	_				
<210> 321					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 321					
tatatatgcg gccgccattc	ccagttcttc	catctg	,		36
<210> 322					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 322					
tatatatgcg gccgcggcca	ttcccagttc	ttccat			36
	_				
<210> 323	•				
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 323					
tatatatgcg gccgcagggg	ccattcccag	ttcttc			36
<210> 324					
<211> 36				•	
<212> DNA					
<213> Synthetic	:				
4400- 224					
<400> 324		COCTTC			36
tatatatgcg gccgcggcag	gggccacce	caytte			50
<210> 325					
::<211> 36					
- <212> DNA					
<213> Synthetic	:		•		
•					
<400> 325					•
tatatatgcg gccgccaggg	caggggccat	tcccag		•	36

<210> 326		
<211> 36		
<212> DNA		
<213> Synthetic		
.400- 206	•	
<400> 326		36
tatatatgcg gccgcctgca gggcaggggc cattcc		20
<210> 327		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 327		م. د
tatatatgcg gccgcgggct gcagggcagg ggccat		36
<210> 328		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 328		20
tatatatgcg gccgcggtgg gctgcagggc aggggc	•	36
·	•	
<210> 329	•	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 329		36
tatatatgcg gccgcctggg tgggctgcag ggcagg		30
<210> 330		
<211> 36		
<212> DNA		
<213> Synthetic		
(213) Synthetic		
<400> 330	•	36
tatatatgcg gccgcaccct gggtgggctg cagggc	•	30
<210> 331		
<211> 36		
<212> DNA		
<213> Synthetic		
12137 Synchectic	•	
.400. 224		
<400> 331		36
tatatatgcg gccgcggcac cctgggtggg ctgcag		30
	· · · · · · · · · · · · · · · · · · ·	
<210> 332	•	
<211> 36		
<212> DNA		
<213> Synthetic		
-400> 222		
<400> 332	•	36
tatatatgcg gccgccatgg caccctgggt gggctg		-
<210> 333		
<211> 36		
<212> DNA		
<213> Synthetic	•	
400× 222		
<400> 333		36
tatatatgcg gccgccggca tggcaccctg ggtggg		J (
·		
<210> 334		
<211> 36	• • • • • • • • • • • • • • • • • • • •	
<212> DNA		
<213> Synthetic	•	

<400> 334	
tatatatgcg gccgcggccg gcatggcacc ctgggt	3.6
<210> 335	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 335	
tatatatgeg geegegaagg eeggeatgge accetg	36
<210> 336	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 336	
	36
tatatatgcg gccgcggcga aggccggcat ggcacc	
<210> 337	
<211> 36	
<212> DNA	
<213> Synthetic	•
<400> 337	2.5
tatatatgeg geegeagagg egaaggeegg catgge	36
<210> 338 <211> 36	
<212> DNA	
<213> Synthetic	
<400> 338	
tatatatgcg gccgcagcag aggcgaaggc cggcat	36
<210> 339	•
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 339	
tatatatgcg gccgcgaaag cagaggcgaa ggccgg	36
<210> 340	
<211> 36	
<212> DNA	
<213> Synthetic	•
<400> 340	
tatatatgcg gccgcctgga aagcagaggc gaaggc	36
<210> 341	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 341	36
tatatatgcg gccgcgct ggaaagcaga ggcgaa	
<210> 342	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 342	31
tatatatgcg gccgcccggc gctggaaagc agaggc	3
<210> 343	
<211> 36	
*# <b>6.6</b> * <b>**</b>	

	<212> DNA					
	<213> Synthetic					
	<400> 343					
tatata	atgeg geegetgeee	ggcgctggaa	agcaga			36
				• 0		
	<210> 344					
	<211> 36					
	<212> DNA					
	<213> Synthetic					
	-				· ·	
	-400> 344					
	<400> 344					
tatat	atgeg geegeteetg	cccggcgctg	gaaagc			36
	<210> 345					
	<211> 36					
	<212> DNA					
	<213> Synthetic					_
	<400> 345					
tatat	atgcg gccgcccctc	ctgcccggcg	ctggaa			36
	<210> 346			•		
	<211> 36					
	<212> DNA				•	
	<213> Synthetic					
	-					
	<400> 346					
						26
tatat	atgcg gccgcgaccc	ctcctgcccg	gegetg			36
-						
	<210> 347					
	<211> 36					
	<212> DNA					
	<213> Synthetic					
	•			′		
	<400> 347					
tatat	atgcg gccgccagga	cccctcctgc	ccggcg			36
	<210> 348					
	<211> 36					
	<212> DNA					
	<213> Synthetic					
	<400> 348					
tarat	atgcg gccgcaacca	ggacccctcc	taccca	•		36
CLCCC	arges googenacea	99444444	-50005			•
	.010- 010					
	<210> 349					
	<211> 36					
	<212> DNA					
	<213> Synthetic					
	-400- 240					
	<400> 349					3.0
tatat	atgcg gccgcagcaa	ccaggacccc	tcctgc			36
	<210> 350					
	<211> 36			•		
	<212> DNA					
	<213> Synthetic					
	<400> 350					
tatat		caaccaggac	ccctcc	*		36
tatat	<400> 350 atgcg gccgcgctag	caaccaggac	ccctcc	•		36
tatat	atgcg gccgcgctag	caaccaggac	ccctcc	•		36
tatat	atgcg gccgcgctag	caaccaggac	ccctcc	·		36
tatat	<pre>.atgcg gccgcgctag      &lt;210&gt; 351      &lt;211&gt; 36</pre>	caaccaggac	ccctcc	·		36
tatat	atgcg gccgcgctag	caaccaggac	ccetce	·		36
tatat	<pre>catgcg gccgcgctag   &lt;210&gt; 351   &lt;211&gt; 36   &lt;212&gt; DNA</pre>		ccctcc	·		36
tatat	<pre>.atgcg gccgcgctag      &lt;210&gt; 351      &lt;211&gt; 36</pre>		ccctcc	·		36
tatat	<pre>catgcg gccgcgctag &lt;210&gt; 351 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>		ccctcc			36
	<pre><atgcg <210="" gccgcgctag=""> 351 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 351</atgcg></pre>					
	<pre>catgcg gccgcgctag &lt;210&gt; 351 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>					36

<210> 352	•	
<211> 36	•	
<212> DNA		
<213> Synthetic		
•		
<400> 352		
tatatatgcg gccgccagat ggctagcaac caggac	3	6
<210> 353		
<211> 36		
<212> DNA		
<213> Synthetic		
_		
<400> 353		
tatatatgcg gccgcctgca gatggctagc aaccag	3	6
	•	
<210> 354		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 354	•	
tatatatgcg gccgcgctct gcagatggct agcaac	·	6
<210> 355		
<211> 36		
<212> DNA		
<213> Synthetic		
·		
<400> 355	_	
tatatatgcg gccgcgaagc tctgcagatg gctagc		16
<210> 356		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 356	4	
tatatatgcg gccgccagga agctctgcag atggct	•	36
	•	
<210> 357		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 357		36
tatatatgcg gccgcctcca ggaagctctg cagatg	•	, 0
244. 254		
<210> 358		
<211> 36		
<212> DNA		
<213> Synthetic		
.400. 250	÷	
<400> 358	•	36
tatatatgcg gccgccacct ccaggaagct ctgcag		
-210> 350		
<210> 359 <211> 36	•	
<211> 36 <212> DNA		
<212> DNA <213> Synthetic		
relar plumerar		
<400> 359		
tatatatgcg gccgccgaca cctccaggaa gctctg		36
raracacycy yeeyeeyaca errecayyaa yeeery		
<210> 360	•	
<210> 360 <211> 36		
<211> 36 <212> DNA	•	
7215 MAN	•	

<400> 360	•		_		
					2.0
tatatatgcg gccgcgtacg	acacctccag	gaagct			3.6
<210> 361					
<211> 36					
<212> DNA					
<213> Synthetic			,		
	•			•	
<400> 361					
-	accacactc	cannaa			36
tatatatgcg gccgcgcggt	acgatatett	cayyaa			30
•					
<210> 362					
<211> 36					
<212> DNA					
<213> Synthetic	•				
<400> 362					
tatatatgcg gccgcaacgc	ggtacgacac	ctccag			36
		.•			
<210> 363					
<211> 36				•	
<212> DNA					
<213> Synthetic				•	
	•				
<400> 363					
tatatatgcg gccgctagaa	cacaatacaa	cacctc			36
tatatatada Ataataa	0909900090				
2.2					
<210> 364					
<211> 36					
<212> DNA					
<213> Synthetic					
4225- 07					
-400> 364					
<400> 364					36
tatatatgcg gccgcgcgta	gaacgcggta	cgacac			36
. 4					
<210> 365			•		
<211> 36				•	
<212> DNA					
<213> Synthetic					
<400> 365					
tatatatgcg gccgcgtggc	gtagaacgcg	gtacga			36
				•	
<210> 366					
	•				
<211> 36					
<212> DNA					
<213> Synthetic	:				
<400> 366					
tatatatgcg gccgcaaggt	ggcgtagaac	gcggta			36
cacacacage googenage		5-35			
-010- 259					
<210> 367					
<211> 36					
<212> DNA					
<213> Synthetic	•				
				•	
<400> 367					
				·	20
tatatatgcg gccgccgcaa	ggtggcgtag	aacgcg			36
<210> 368					
<211> 39	•				
<212> DNA					
<213> Synthetic					
-211/ Synthetic	•				
			•		
<400> 368					
tatatatoro occorcioco	caangtagco	TAGAACGCG			39

## INTERNATIONAL SEARCH REPORT

nel Application No PCT/US 99/20891

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/10 C12N C12N15/25 C12N15/62 C07K14/535 C07K14/005 According to International Patent Classification (IPC) or to both national classification and IPC **B. FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) C12N C07K IPC 7 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. X MCWHERTER C (REPRINT) ET AL: "Protein 14,15 engineering of the myelopoietins by mutation, circular permutation and combination of amino acid sequences." BLOOD, (15 NOV 1997) VOL. 90, NO. 10, PART 2, SUPP. '1!, PP. 3531-3531. PUBLISHER: W B SAUNDERS CO, INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA, PA 19106-3399. , XP000876993 Abstract no. 3531 Y 1-13 abstract Further documents are listed in the continuation of box C. Petent family members are fieled in armex. X Special categories of cited documents: "I" later document published after the international filing date or priority date and not in conflict with the application but clied to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone. "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of snother clation or other special reason (as specified) "\" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such docu-ments, such combination being obvious to a person skilled "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 24 February 2000 14/03/2000 **Authorized officer** Name and making address of the ISA European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rignells Tel. (+31-70) 340-2040, Tx. 91 651 epo ni, Fex: (+91-70) 340-3016 Hornig, H

Form PCT/ISA/210 (second sheet) (July 1992)

1

## INTERNATIONAL SEARCH REPORT

Intern. nal Application No PCT/US 99/20891

		FC1/03 99/20091
	ntion) DOCUMENTS CONSIDERED TO BE RELEVANT	Relevant to claim No.
ategory *	Citation of document, with indication, where appropriate, of the relevant passages	Piorovala to Calain Nec
(	MCKEARN J (REPRINT) ET AL: "Structure and function of myelopoietins, a family of hematopoietic growth factors." BLOOD, (15 NOV 1997) VOL. 90, NO. 10, PART 1, SUPP. '1!, PP. 249-249. PUBLISHER: W B SAUNDERS CO, INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA, PA 19106-3399. , XP000876981	14,15
Υ .	Abstract no .249 abstract	1-13
X	WO 97 12977 A (SEARLE & CO ; ZURFLUH LINDA L (US); KLEIN BARBARA K (US); MCWHERTER) 10 April 1997 (1997-04-10)	14
Y	the whole document	1-13
X	S. MERLIN ET AL.: "Phage presentation and affinity-selection of a deletion mutant of human interleukin-3" APPLIED BIOCHEMISTRY AND BIOTECHNOLOGY, vol. 67, no. 3, September 1997 (1997-09), pages 199-214, XP000884149 HUMANA PRESS INC., CLIFTON, N.J., US	14
Y	the whole document	1-13
Υ .	H. GRAM ET AL.: "Phage display as a rapid gene expression system: production of bloactive cytokine-phage and generation of neutralizing monoclonal antibodies" J. IMMUNOL. METHODS, vol. 161, 1993, pages 169-176, XP002131429 ELSEVIER, AMSTERDAM NL the whole document	1-13
<b>Y</b>	CLACKSON T ET AL: "IN VITRO SELECTION FROM PROTEIN AND PEPTIDE LIBRARIES" TIBTECH, GB, CAMBRIDGE, vol. 12, 1 May 1994 (1994-05-01), pages 173-184, XP000652419 cited in the application the whole document	1–13
X	R. GRAF AND H.K. SCHACHMAN: "Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase" PROC. NATL. ACAD. SCI., vol. 93, October 1996 (1996—10), pages 11591—11596, XP002131430 NATL. ACAD. SCI., WASHINGTON, DC, US; cited in the application	14
Ÿ	the whole document	1-13

Inters. And Application No PCT/US 99/20891

(Continu	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	· · · · · · · · · · · · · · · · · · ·
cheboth .	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	HORLICK R A ET AL: "PERMUTEINS OF INTERLEUKIN 1 BETA-A SIMPLIFIED APPROACH FOR THE CONSTRUCTION OF PERMUTATED PROTEINS HAVING NEW TERMINI" PROTEIN ENGINEERING, GB, OXFORD UNIVERSITY PRESS, SURREY, vol. 5, no. 5, 1 January 1992 (1992-01-01), pages 427-431, XP002022097	14
	ISSN: 0269-2139 cited in the application the whole document	
X	KREITMAN R J ET AL: "A CIRCULARLY PERMUTED RECOMBINANT INTERLEUKIN 4 TOXIN WITH INCREASED ACTIVITY" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, US, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 91, no. 15, 1 July 1994 (1994-07-01), pages 6889-6893, XP002022099 ISSN: 0027-8424 the whole document	14
P,X	MCWHERTER C A ET AL: "Circular permutation of the granulocyte colony-stimulating factor receptor agonist domain of myelopotetin." BIOCHEMISTRY, (1999 APR 6) 38 (14) 4564-71. JOURNAL CODE: AOG., XP002131431 the whole document	14,15
T	T.J. MACVITTIE ET AL.: "Myelopoietin, an engineered chimeric IL-3 and G-CSF receptor agonist, stimulates multilineage hematopoietic recovery in a nonhuman model of radiation-induced myelosuppression" BLOOD, vol. 95, no. 3, 1 February 2000 (2000-02-01), pages	
	837-845, XP002131432 SAUNDERS, DULUTH, NEW YORK, US the whole document	

Form PCT/18A/2:10 (continuation of second sheet) (July 1992)

information on patent family members

PCT/US 99/20891

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9712977 A	10-04-1997	AU 7390096 A CA 2234042 A EP 0859843 A	28-04-1997 10-04-1997 26-08-1998

```
<213> Synthetic
      <400> 317
                                                                        36
tatatatgcg gccgcttcca tctgctgcca gatggt
     <210> 318
     <211> 36
      <212> DNA
      <213> Synthetic
     <400> 318
                                                                        36
tatatatgcg gccgcttctt ccatctgctg ccagat
      <210> 319
     <211> 36
      <212> DNA
     <213> Synthetic
      <400> 319
                                                                        36
tatatatgcg gccgccagtt cttccatctg ctgcca
      <210> 320
      <211> 36
      <212> DNA
      <213> Synthetic
      <400> 320
                                                                         36
tatatatgcg gccgctccca gttcttccat ctgctg
      <210> 321
      <211> 36
      <212> DNA
      <213> Synthetic
      <400> 321
                                                                         36
tatatatgcg gccgccattc ccagttcttc catctg
       <210> 322
       <211> 36
       <212> DNA
       <213> Synthetic
      <400> 322
                                                                         36
tatatatgcg gccgcggcca ttcccagttc ttccat
       <210> 323
       <211> 36
       <212> DNA
       <213> Synthetic
       <400> 323
                                                                         36
 tatatatgcg gccgcagggg ccattcccag ttcttc
       <210> 324
       <211> 36
       <212> DNA
       <213> Synthetic
       <400> 324
                                                                         36
 tatatatgcg gccgcggcag gggccattcc cagttc
       <210> 325
       <211> 36
       <212> DNA
       <213> Synthetic
       <400> 325
                                                                         36
 tatatatgcg gccgccaggg caggggccat tcccag
```

<210> 326	_				
<211> 36			•	•	
<212> DNA					
<213> Synthetic					
2213> Synthetic					
<400> 326				•	36
tatatatgcg gccgcctgca	gggcaggggc (	cattcc			30
<210> 327	•				
<211> 36				•	
<212> DNA					
<213> Synthetic					
. (213) Synthetic					
<400> 327		·	•		36
tatatatgcg gccgcgggct	gcagggcagg	ggccat			30
			•		
<210> 328					
<211> 36				•	
<212> DNA					
<213> Synthetic					•
			•		
<400> 328		•		•-	
	setsessess	=00000			3.6
tatatatgcg gccgcggtgg	gergeaggge	299990			
<210> 329					
<211> 36					
<212> DNA					
<213> Synthetic					
				·	
<400> 329					
tatatatgcg gccgcctggg	tagactacaa	ggcagg			36
tutututgeg geograps	-33355			• •	
<210> 330	•				
			•		
<211> 36	•				
<212> DNA					•
<213> Synthetic			•		
				•	
<400> 330					26
tatatatgcg gccgcaccct	gggtgggctg	cagggc		•	36
<210> 331					
<211> 36					
<212> DNA					
<213> Synthetic					
7215° 07.1010121					
<400> 331		•			
		atacas	9.0		36
tatatatgcg gccgcggcac	cccgggcggg	cegeag			-
			,		
<210> 332			·		
<211> 36					
<212> DNA					
<213> Synthetic					
•				•	
<400> 332					
tatatatgcg gccgccatgg	caccctgggt	gggctg	•		36
				•	
<210> 333				•	
			•	•	
<211> 36				•	
<212> DNA					
<213> Synthetic				•	
		-			
<400> 333					
tatatatgcg gccgccggca	tggcaccctg	ggtggg			36
		-			•
<210> 334	_	•	•	•	
<211> 36	•				
<211> 36 <212> DNA					
SZIZ> LKNA					
<213> Synthetic					

<400> 334 tatatatgcg gccgcggccg gcatggcacc ctgggt -	36
<210> 335	
<211> 36	
<212> DNA	
<213> Synthetic	
400> 225	
<400> 335 tatatatgcg geogegaagg coggoatggc accetg	36
<210> 336	
<211> 36	•
<212> DNA <213> Synthetic	
<400> 336	20
tatatatgeg geegeggega aggeeggeat ggeace	36
<210> 337	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 337	Ò
tatatatgcg gccgcagagg cgaaggccgg catggc	36
<210> 338	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 338	0.2
tatatatgcg gccgcagcag aggcgaaggc cggcat	36
<210> 339	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 339	
tatatatgcg gccgcgaaag cagaggcgaa ggccgg	36
<210> 340	•
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 340	
tatatatgeg geegeetgga aageagagge gaagge	36
<210> 341	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 341	
tatatatgcg gccgcgctc ggaaagcaga ggcgaa	36
<210> 342	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 342	
tatatatgcg gccgcccggc gctggaaagc agaggc	36
<210> 343	
<211> 36	

<212> DNA <213> Synthetic				-	
<400> 343					
tatatatgcg gccgctgccc	ggcgctggaa	agcaga			36
<210> 344					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 344					36
tatatatgcg gccgctcctg	eceggegetg	gaaagc			36
<210> 345					
<211> 36					
<212> DNA					
<213> Synthetic					
_					
<400> 345	•				
tatatatgcg gccgcccctc	ctgcccggcg	ctggaa			. 36
<210> 346				•	
<211> 36					•
<212> DNA					
<213> Synthetic					
				•	
<400> 346					2.0
tatatatgcg gccgcgaccc	ctcctgcccg	gegetg			36
<210> 347					
<211> 36					
<212> DNA					
<213> Synthetic					
_					
<400> 347		000000			36
tatatatgcg gccgccagga	ceeeteetge	eeggeg			30
<210> 348					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 348					
tatatatgcg gccgcaacca	ggacccctcc	tgeceg			36
-210- 240					
<210> 349		•			
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 349					
tatatatgcg gccgcagcaa	ccaggacccc	tcctgc			36
<210> 350					
<211> 36					
<211> JU <212> DNA					
	•		i		
<213> Synthetic					
<400> 350					
tatatatgcg gccgcgctag	caaccaggac	ccctcc	·		36
<210> 351					
<211> 36					•
<212> DNA					
<213> Synthetic					
<400> 351				8	
tatatatgcg gccgcatggc	tagcaaccag	gacccc			36

<210> 352	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 352	36
tatatatgcg gccgccagat ggctagcaac caggac	
<210> 353	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 353	76
tatatatgcg gccgcctgca gatggctagc aaccag	36
<210> 354	
<211> 36	
<211> 36 <212> DNA	
<213> Synthetic	
(213) Symmetra	
<400> 354	
tatatatgcg gccgcgctct gcagatggct agcaac	36
<210> 355	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 355	36
tatatatgcg gccgcgaagc tctgcagatg gctagc	
<210> 356	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 356	
tatatatgcg gccgccagga agctctgcag atggct	36
<210> 357	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 357	
tatatatgcg gccgcctcca ggaagctctg cagatg	36
<210> 358	
<211> 36	
<212> DNA	
<213> Synthetic	
·	
. <400> 358	36
tatatatgcg gccgccacct ccaggaagct ctgcag	30
<210> 359	
<211> 36	
<211> 30 <212> DNA	
<213> Synthetic	
<400> 359	
tatatatgcg gccgccgaca cctccaggaa gctctg	36
<210> 360	
<211> 36	
<212> DNA	
<213> Synthetic	

<400> 360 tatatatgcg gccgcgtacg	acacetecaq	gaaget		-	36
tatatatgeg geegestacg	deacceeag	gg			
<210> 361					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 361			•		
tatatatgcg gccgcgcggt	acgacacctc	caggaa			36
<210> 362					
<211> 36					
<212> DNA					
<213> Synthetic				•	
<400> 362					
tatatatgcg gccgcaacgc	ggtacgacac	ctccag			36
<210> 363					
<211> 36				•	
<212> DNA					
<213> Synthetic	<b>:</b> .			•	
<400> 363					36
tatatatgcg gccgctagaa	egeggtacga	Eacete			
<210> 364					
<211> 36		•			
<212> DNA					
<213> Synthetic	:				
<400> 364 tatatatgcg gccgcgcgta	gaacgcggta	cgacac			36
tatatatycy getycycyc	. gaacgeggea	Oguduo			
<210> 365					
<211> 36					
<212> DNA					
<213> Synthetic	:				
<400> 365			•		
tatatatgcg gccgcgtgg	gtagaacgcg	gtacga			36
tututudgag gaagagaga	<b>3</b> ,-3	3			
<210> 366	·				
<211> 36					
<212> DNA					
<213> Synthetic	c				
<400> 366					
tatatatgcg gccgcaagg	t ggcgtagaac	gcggta			36
<210> 367					
<211> 36					
<212> DNA	_				
<213> Synthetic	C				
<400> 367			ł		
tatatatgcg gccgccgca	a ggtggcgtag	aacgcg			36
<210> 368					
<211> 39	-				
<211> 39					
<213> Syntheti	c				
<400> 368					39
		ransarnro			

Inter: nel Application No PCT/US 99/20891

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/10 C12I C07K14/535 C07K14/005 C12N15/25 C12N15/62 According to international Patent Classification (IPC) or to both national classification and IPC B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. Category \* 14,15 MCWHERTER C (REPRINT) ET AL: "Protein X engineering of the myelopoietins by mutation, circular permutation and combination of amino acid sequences."
BLOOD, (15 NOV 1997) VOL. 90, NO. 10, PART
2, SUPP. '1!, PP. 3531-3531. PUBLISHER: W B SAUNDERS CO, INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA, PA 19106-3399., XP000876993 1-13 Y Abstract no. 3531 abstract Patent family members are listed in annex. Further documents are listed in the continuation of box C. X X Special categories of cited documents: "I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the International "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone filing date "L" document which may throw doubts on priority claim(s) or which is clied to establish the publication date of another citation or other special reason (as specified) "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such document. "O" document referring to an oral disclosure, use, exhibition or ments, such combination being obvious to a person sidled document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 14/03/2000 24 February 2000 Name and making address of the ISA **Authorized officer** European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo ni, Hornig, H Fex: (491-70) 940-9016

001000EA1 | -

Intern and Application No PCT/US 99/20891

		C1/US 99/20891
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	Relevant to claim No.
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Protovala to case i vec
X	MCKEARN J (REPRINT) ET AL: "Structure and function of myelopoietins, a family of hematopoietic growth factors." BLOOD, (15 NOV 1997) VOL. 90, NO. 10, PART 1, SUPP. '1!, PP. 249-249. PUBLISHER: W B SAUNDERS CO, INDEPENDENCE SQUARE WEST CURTIS CENTER, STE 300, PHILADELPHIA, PA	14,15
Y	19106-3399. , XP000876981 Abstract no .249 abstract	1-13
X	WO 97 12977 A (SEARLE & CO ; ZURFLUH LINDA L (US); KLEIN BARBARA K (US); MCWHERTER)	14
Y	10 April 1997 (1997-04-10) the whole document	1-13
X	S. MERLIN ET AL.: "Phage presentation and affinity-selection of a deletion mutant of human interleukin-3" APPLIED BIOCHEMISTRY AND BIOTECHNOLOGY, vol. 67, no. 3, September 1997 (1997-09), pages 199-214, XP000884149 HUMANA PRESS INC., CLIFTON, N.J., US	14
Υ.	the whole document	1-13
Y	H. GRAM ET AL.: "Phage display as a rapid gene expression system: production of bloactive cytokine-phage and generation of neutralizing monoclonal antibodies" J. IMMUNOL. METHODS, vol. 161, 1993, pages 169-176, XP002131429 ELSEVIER, AMSTERDAM NL the whole document	1-13
Y	CLACKSON T ET AL: "IN VITRO SELECTION FROM PROTEIN AND PEPTIDE LIBRARIES" TIBTECH, GB, CAMBRIDGE, vol. 12, 1 May 1994 (1994-05-01), pages 173-184, XP000652419 cited in the application the whole document	1-13
X	R. GRAF AND H.K. SCHACHMAN: "Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase" PROC. NATL. ACAD. SCI., vol. 93, October 1996 (1996-10), pages 11591-11596, XP002131430 NATL. ACAD. SCI., WASHINGTON, DC, US;	14
Y	cited in the application the whole document	1–13

Intern. .nel Application No PCT/US 99/20891

		PC1/US 99/20891
C.(Continue	ition) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HORLICK R A ET AL: "PERMUTEINS OF INTERLEUKIN 1 BETA-A SIMPLIFIED APPROACH FOR THE CONSTRUCTION OF PERMUTATED PROTEINS HAVING NEW TERMINI" PROTEIN ENGINEERING, GB, OXFORD UNIVERSITY PRESS, SURREY, vol. 5, no. 5, 1 January 1992 (1992-01-01), pages 427-431, XP002022097 ISSN: 0269-2139 cited in the application the whole document	14
X	KREITMAN R J ET AL: "A CIRCULARLY PERMUTED RECOMBINANT INTERLEUKIN 4 TOXIN WITH INCREASED ACTIVITY" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 91, no. 15, 1 July 1994 (1994-07-01), pages 6889-6893, XP002022099 ISSN: 0027-8424 the whole document	14
P,X	MCWHERTER C A ET AL: "Circular permutation of the granulocyte colony-stimulating factor receptor agonist domain of myelopoietin." BIOCHEMISTRY, (1999 APR 6) 38 (14) 4564-71. JOURNAL CODE: AOG., XPO02131431 the whole document	14,15
T	T.J. MACVITTIE ET AL.: "Myelopoietin, an engineered chimeric IL-3 and G-CSF receptor agonist, stimulates multilineage hematopoietic recovery in a nonhuman model of radiation-induced myelosuppression" BLOOD, vol. 95, no. 3, 1 February 2000 (2000-02-01), pages 837-845, XPO02131432 SAUNDERS, DULUTH, NEW YORK, US the whole document	
•		

1

information on patent family members

Intern. And Application No PCT/US 99/20891

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9712977 A	10-04-1997	AU 7390096 A CA 2234042 A EP 0859843 A	28-04-1997 10-04-1997 26-08-1998

# WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7:		(11) International Publication Number: WO 00/18905
C12N 15/10, 15/25, 15/62, C07K 14/535, 14/005	A1	(43) International Publication Date: 6 April 2000 (06.04.00)
(21) International Application Number: PCT/US (22) International Filing Date: 24 September 1999 (		BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD,
(30) Priority Data: 60/101,908 25 September 1998 (25.09.5) (71) Applicant (for all designated States except US): G.D. & CO. [US/US]; Corporate Patent Department, 5110, Chicago, IL 60680-5110 (US).	SEARI	
<ul> <li>(72) Inventor; and</li> <li>(75) Inventor/Applicant (for US only): LEE, Stephen, C. 828 Fernview Drive, Creve Coeur, MO 63141 (U</li> <li>(74) Agents: BAUER, S., Christopher et al.; G.D. Co., Corporate Patent Dept., P.O. Box 5110, Ct 60680-5110 (US).</li> </ul>	S). Searle	With international search report.  Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of
(54) Title: METHOD OF PRODUCING PERMUTEINS		
linker 1 174 1 17  hG-CSF gene hG-CSF gene hG-CSF Tandem repear precursor p	4	FOR INDIVIDUAL PCRIAMPLIFICATIONS  POOL 174 1 PCR PRODUCTS, INSERT INTO  DISPLAY VECTOR, 174 1 173 MAKE PHAGEMID 1 174
		PARTICLES COLLECTION OF INDIVIDUAL

(57) Abstract

A method of producing circularly-permuted proteins (permuteins) by scanning permutagenesis comprises making and inserting a series of circularly-permuted genes into a display vector, expressing these genes such that the gene products are localized to the surface of the display vector, generating a library of display vectors presenting the permuted protein, affinity-selecting the display vectors with a target protein that can bind the permuted protein, isolating and analyzing clones of selected display vectors to identify the circularly-permuted protein. The invention further discloses methods of expressing and uses of permuteins.

Library of permuted proteins presented on phage

PERMUTED GENES

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	Prance	LU	Luxembourg	. SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo .
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	<b>Greece</b>		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali .	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	` UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	. NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	·KP	Democratic People's	NZ	New Zealand		
СМ	Cameroon		Republic of Korea	PL	Poland		
CN	China .	KR	Republic of Korea	PT	Portugal		
CU	Cuba	. KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	L	Liechtenstein	SD	Sudan		•
DK	Denmark	LK	Sri Lanka	SE	Sweden		• .
EE	Estonia	LR	Liberia	SG	Singapore		•

WO 00/18905 PCT/US99/20891

# Method of producing permuteins by scanning permutagenesis

#### **Priority**

The present application claims priority under Title 35, United States Code, § 119 of United States Provisional Application Serial No. 60/101,908, filed September 25, 1998.

### Field of the invention

A method of producing circularly-permuted proteins (permuteins) by scanning permutagenesis comprises making and inserting a series of circularly-permuted genes into a display vector, expressing these genes such that the gene products are localized to the surface of the display vector, generating a library of display vectors presenting the permuted protein, affinity-selecting the display vectors with a target protein that can bind the permuted protein, isolating and analyzing clones of selected display vectors to identify the circularly-permuted protein. The invention further discloses methods of expressing and uses of permuteins.

## Background of the invention

#### Protein permutagenesis

Circularly permuted proteins are made by reordering the primary sequence of a parent protein. The amino and carboxy terminal ends of the parent protein are joined by a peptide linker and new amino and carboxy terminal ends are generated at other positions in the sequence. This technique of generating variants has been applied to a wide variety of proteins (Table 1).

Circularly permuted proteins, in many cases, are structurally and functionally similar to their non-permuted parent molecule after they undergo refolding. The information necessary to direct the folding of proteins into tertiary structures is present in secondary structural domains. Vectorial folding of proteins from their native amino to carboxy ends is not often observed. The ability of permuteins to retain structural and functional properties is remarkable, extending earlier observations on the plasticity of proteins with respect to amino acid

30

5

10

15

20

25

10

15

20

25

30

35

substitutions (Olins P.O. et al., J. Biol. Chem. 270: 23754-23760, 1995; Lowman and Wells, J. Mol. Biol. 234: 564-578, 1993) and short amino acid insertions (Sondek, J. and D. Shortle, Proteins 7: 387-393, 1990; Shortle, D. and J. Sondek, Curr. Opin. Biotechnol. 6: 299-305).

### Protein sequence reorganization

Rearrangements of DNA sequences serve an important role in evolution by generating a diversity of new proteins differing in structure and function. Gene duplication and exon shuffling, for example, generate diversity and provide organisms with a competitive advantage since the basal mutation rate is low (Doolittle, *Protein Science* 1: 191-200, 1992).

Recombinant DNA methods have facilitated studies on the effect of sequence transposition on protein folding, structure, and function. rearrangement of proteins using this approach was described by Goldenberg and Creighton (J. Mol. Biol. 165:407-413, 1983). A new N-terminus is selected at an internal site (breakpoint) of the original sequence, the new sequence having the same order of amino acids as the original from the breakpoint until it reaches an amino acid that is at or near the original C-terminus. At this point the new sequence is joined, either directly or through an additional portion of sequence (linker), to an amino acid that is at or near the original N-terminus, and the new sequence continues with the same sequence as the original until it reaches a point that is at or near the amino acid that was N-terminal to the breakpoint site of the original sequence, this residue forming the new C-terminus of the chain. Similar approaches have also been used in other studies (Cunningham et al., Proc. Natl. Acad. Sci. U.S.A. 76:3218-3222, 1979; Teather & Erfle, J. Bacteriol. 172: 3837-3841, 1990; Schimming et al., Eur. J. Biochem. 204: 13-19, 1992; Yamiuchi and Minamikawa, FEBS Lett. 260:127-130, 1991: MacGregor et al., FEBS Lett. **378**:263-266, 1996).

These general approaches have been applied to proteins which range in size from 58 to 462 amino acids (Goldenberg & Creighton, J. Mol. Biol. 165:407-413, 1983; Li & Coffino, Mol. Cell. Biol. 13:2377-2383, 1993). The proteins represent a broad range of structural classes, including proteins that contain predominantly alpha helix (interleukin-4; Kreitman et al., Cytokine 7:311-318, 1995), beta sheet (interleukin-1; Horlick et al., Protein Eng. 5:427-431, 1992), or mixtures of the two types of secondary structures (yeast phosphoribosyl anthranilate isomerase; Luger et al., Science 243:206-210, 1989).

10

15

20

25

30

35

Although broad categories of protein function are represented in these sequence reorganization studies, the results of these studies have been highly variable. In many cases substantially lower activity, solubility, or thermodynamic observed coli dihydrofolate reductase. were Œ. stability transcarbamoylase, phosphoribosyl anthranilate isomerase, glyceraldehyde-3phosphate dehydrogenase, ornithine decarboxylase, ompA, yeast phosphoglycerate dehydrogenase). In other cases, the sequence rearranged protein appeared to have many nearly identical properties as its natural counterpart (basic pancreatic trypsin inhibitor, T4 lysozyme, ribonuclease T1, Bacillus β-glucanase, interleukin-1β, α-spectrin SH3 domain, pepsinogen, interleukin-4). In exceptional cases, an unexpected improvement over some properties of the natural sequence was observed, e.g., the solubility and refolding rate for rearranged a-spectrin SH3 domain sequences, and the receptor affinity and anti-tumor activity of transposed interleukin-4-Pseudomonas exotoxin fusion molecule (Kreitman et al., Proc. Natl. Acad. Sci. U.S.A. 91:6889-6893, 1994; Kreitman et al., Cancer Res. 55:3357-3363. 1995).

The primary motivation for reorganization studies has been to study the role of short-range and long-range interactions in protein folding and stability. Sequence rearrangements of this type convert a subset of interactions that are long-range in the original sequence into short-range interactions in the new sequence, and vice versa. The fact that many of these sequence rearrangements are able to attain a conformation with at least some activity is persuasive evidence that protein folding occurs by multiple folding pathways (Viguera et al., J. Mol. Biol. 247:670-681, 1995). In the case of the SH3 domain of alpha-spectrin, choosing new termini at locations that corresponded to beta hairpin turns resulted in proteins with slightly less stability, but which were nevertheless able to fold.

The positions of the internal breakpoints used in the studies cited above are found exclusively on the surface of proteins, and are distributed throughout the linear sequence without any obvious bias towards the ends or the middle (the variation in the relative distance from the original N-terminus to the breakpoint is ca. 10 to 80% of the total sequence length). The linkers connecting the original N-and C-termini in these studies have ranged from 0 to 9 residues. In one case (Yang & Schachman, Proc. Natl. Acad. Sci. U.S.A. 90:11980-11984, 1993), a portion of sequence has been deleted from the original C-terminal segment, and the connection made from the truncated C-terminus to the original N-terminus. Flexible hydrophilic residues such as Gly and Ser are frequently used in the linkers. Viguera et al.(J. Mol. Biol. 247:670-681, 1995) compared joining the

original N- and C- termini with 3- or 4-residue linkers; the 3-residue linker was less thermodynamically stable. Protasova et al. (*Protein Eng.* 7:1373-1377, 1994) used 3- or 5-residue linkers in connecting the original N-termini of *E. coli* dihydrofolate reductase; only the 3-residue linker produced protein in good yield.

5

10

Protein permutagenesis can be used to optimize the activity of fusion proteins or proteins conjugated to other molecules. A fusion between interleukin-4 (IL-4) and *Pseudomonas* exotoxin has been permuted resulting in a protein that has the first amino acid of the IL-4 domain at position 38 and the new carboxy end occurs at amino acid position 37 (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994). The permuted fusion has increased affinity for the IL-4 receptor, increased cytotoxicity to IL-4 receptor bearing renal carcinoma cells, and increased anti-tumor activity in a murine model, compared to the non-permuted parent fusion protein (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* 55:3357-3363, 1995; Puri, R. K. et al., *Cellular Immunol.* 171: 80-86, 1996). Increased potency of the permuted molecule is believed to result from a reduction in steric interference between the IL-4 domain in the parent molecule and its receptor.

15

20

Steric hindrance is likely to be a concern for other chimeric proteins which interact with receptors through a relatively large area of their surface. The same issue also arises with bioconjugates, containing relatively small chemicals conjugated to proteins or other molecules in complex polymers (Rose, K. et al., *Molecular Immunology* 32: 1031-1037, 1995).

## Phage display methods

25

30

Display methods allow affinity selection of protein variants from a library of displayed proteins or peptides (Clackson, T. and J.A. Wells, Tibtech 12: 173-184, 1994; Winter, G., Drug Development Res. 33: 71-89, 1994). Many biological entities can be used in display methodologies (so-called "genetic packages" for presentation, including bacterial and eukaryotic cells, various eukaryotic and prokaryotic viruses, and spores), but the most commonly used vehicles used for display are filamentous bacteriophage, as used herein. We envision the possibility that a genetic package other than the particular phage used here could be used to present libraries of permuteins, and if so, constitute essentially the same invention.

Foreign proteins are presented on the surface of a phage particle, and the gene encoding the foreign protein is encapsulated in the virion. Because they are linked by the phage particle, affinity isolation of the presented protein also leads to affinity isolation of the corresponding genes. Extremely large libraries of phage presented proteins are constructed and affinity screened very rapidly. From the standpoint of how quickly mutant proteins can be made and screened for activity, phage display is the most efficient mutagenesis technique currently available.

### Functional properties of permuteins

Permuteins can have improved biological properties by acting through several mechanisms. The permutein acting on the same type of cell as its parent molecule, may have increased binding, or other action, by virtue of increased avidity. Dimers or higher order multimers of these proteins with themselves or other chemical groups, including proteins, can have increased efficacy or potency, or both.

15

10

5

Permuteins can also have improved therapeutic properties through a variety of mechanisms such as: (1) alterations in the overall on- or off-rates or K, or K, of the ligand(s) on the target cell; (2) activation or blockade of complementary receptor signaling pathways; and/or (3) more specific targeting of to the cell of interest. The permuteins may also possess a unique pharmacokinetic distribution and clearance profile (Dehmer et al., Circulation, 91, 2188-2194, 1995; Tanaka et al., Nature Medicine, 3, 437-442, 1997).

20

Permuteins can also have improved properties in vivo, compared to the two components individually, as a result of alterations in biodistribution or half-life. The improved properties can also result from the binding of the permutein to one or more of the receptors, pharmacokinetics, or uptake of the permutein is altered in a favorable manner.

25

Molecular biology approaches have traditionally been used to permute proteins (Horlick, R.A. et al., *Protein Engineering* 5: 427-431, 1992) although chemical approaches have been used to make small permuted proteins (Goldenberg, D. P. and T. E. Creighton, *J. Mol. Biol.* 165: 407-413, 1983). These approaches are relatively labor intensive, limiting the number of permuteins that can be generated and efficiently screened for the desired biological activities. Rapid methods of generating permuteins, coupled with efficient methods for screening are needed that will result in the identification of novel active molecules.

30

## Summary of the invention

The present invention is an improved method for generating permuteins (scanning permutagenesis) based on the display of proteins on bacteriophage surface proteins. Phage display is a powerful, yet convenient tool, traditionally used for mutagenesis and screening (Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994). Improvements to this technology allow the rapid generation and screening of libraries of permuteins. Variables, such as position of the new termini and the length and composition of peptide linkers can easily be varied to generate libraries of the desired diversity.

10

15

5

The present invention relates to methods of producing biologically-active circularly permuted proteins of the formula C¹-L¹-N¹, derived from a parent protein of the formula N¹-C¹, wherein C¹ is comprised of a segment derived from the carboxy portion of said parent protein; N¹ is comprised of a segment derived from the amino terminal portion of said parent protein; and L¹ is a chemical bond or a linker, linking C¹ to the amino terminus of L¹ and carboxy terminus of L¹ to the amino terminus of N¹; comprising the steps of: (a) making a series of circularly-permuted genes; (b) inserting said circularly-permuted genes into a display vector; (c) expressing said circularly-permuted genes such that the proteins encoded by said genes are presented on the surface of the display vector; (d) generating a library of display vectors presenting the expressed circularly permuted protein; (e) affinity-select the presenting display vectors with a target protein that can bind a biologically-active circularly-permuted protein; (f) isolate and analyze clones of selected display vectors to identify the presented circularly-permuted protein.

20

Preferably the method of making a series of circularly-permuted genes is selected from the group consisting of making a tandemly-repeated intermediate, total synthesis of a synthetic gene, assembly of a gene from synthetic oligonucleotides, DNA amplification, and limited digestion of a circular intermediate.

30

25

Preferably, the display vector is selected from the group consisting of bacteriophage display vectors, bacteria, and baculovirus vectors. Even more preferably the presentation vector is a bacteriophage. Even more preferably, the presentation vector is bacteriophage M13. Most preferably, the presentation vector is a bacteriophage M13 gene III vector.

10

15

20

25

30

Preferably the method of making a series of circularly permuted genes is a method of making a tandem repeat intermediate. Even more preferably circularly permuted genes are amplified from the repeat by gene amplification.

Preferably the method of affinity selection comprises the steps consisting of
(a) binding said presentation display vectors to a target protein; (b) eluting said
display vectors; (c) amplifying said display vectors; and (d) biopanning a pool of
said amplified display vectors.

Preferably, the length of  $C^1$  in the permutein is longer than the length of  $C^1$  in said parent protein. More preferably, the length of  $C^1$  in the permutein is shorter than the length of  $C^1$  in said parent protein. Most preferably, the length of  $C^1$  in the permutein is the same length as the length of  $C^1$  in said parent protein.

Preferably, the length of  $N^1$  in the permutein is longer than the length of  $N^1$  in said parent protein. More preferably, the length of  $N^1$  in the permutein is shorter than the length of  $N^1$  in said parent protein. Most preferably, the length of  $N^1$  in the permutein is the same length as the length of  $N^1$  in said parent protein.

The invention also contemplates circularly permuted proteins of the formula C'-L'-N' made by the method of scanning permutagenesis. Preferably, the DNA sequence encoding said linker L' is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 368.

Preferably, the circularly-permuted protein is the G-CSF receptor agonist domain of a species of mylepoietin (MPO). MPO is one member of a family of novel dual cytokine receptor agonists (McKearn, J.P., Myelorestorative activities of synthokine and myelopoietin. In Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996) which are amenable to manipulation by phage display (Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997; Lee, S.C., Phage presentation for cytokine engineering. IBC's Second International Conference on Display Technologies, 1997).

## Brief description of the figures

## Figure 1. Schematic depiction of scanning permutagenesis

Plate A of Figure 1 shows the strategy to generate a scanning permutagenesis phage display library. A plasmid containing directly-repeated

tandem copies of the hG-CSF gene, for example, is constructed by standard methods. The tandem repeat plasmid is used as the template for PCR amplification of genes encoding permuted proteins. Each copy of the G-CSF gene is indicated in light gray (turquoise), and a DNA segment encoding a peptide linker is indicated in dark gray (red).

In individual PCR reactions, oligonucleotide primers that initiate PCR polymerization at the first nucleotide of a chosen codon of G-CSF, and directing polymerization to the end of the tandem construct specifying the carboxy end of the protein encoded on the template is annealed to the tandem template. Also, a second specific primer is also annealed to the template that initiates polymerization at the last nucleotide of the codon encoding the amino acid immediately preceding the codon where polymerization begins with first primer, and which directs polymerization in the opposite direction from that first primer. Amplification between these two primers produces a DNA segment encoding a permuted protein. For example, amplification between the primer indicated by a black arrow initiating at codon 2 and the primer indicated by the blue arrow and initiating at the codon before 2 (codon 1) produces an amplified gene encoding a permuted protein whose amino terminal residue is amino acid 2 of the native protein, and whose final amino acid is amino acid 1 of the native protein.

20

25

5

10

15

A linker peptide is present between the first and final amino acids of the parent protein (residues 1 and 174 in this example). A total of 174 individual amplifications would produce a complete collection of all permuted proteins of this example. More limited collections containing only a selected set of permuteins can be made, as well as more extensive collections made from multiple tandem template plasmids, each containing a different linker sequence between the first and last residues of the two directly repeated tandem gene sequences. The collection of amplified segments can then be inserted into a phagemid presentation vector by standard methods. Phagemid particles produced from these presentation constructs are the scanning permutagenesis phage display library.

30

35

Plate B of Figure 1 shows the affinity screening of a phage display library (See Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994; Winter, G., *Drug Development Res.* 33: 71-89, 1994). In this example, a hG-CSF scanning permutagenesis library as described in Figure 1A is screened using the hG-CSF receptor expressed on mammalian cells as the affinity reagent. In Figure 1B, individual presented proteins are indicated by the shaded circles or diamonds and the affinity reagent is indicated by the light gray (pink) rectangles. Presentation

-10

15

20

25

30

library particles are exposed to affinity reagent, unbound particles are washed away, and receptor-bound particles are eluted. The eluted particles are amplified in *E. coli*, and the screening cycle is repeated. During any round of the screening cycle, the genes encoded (in the present example encoding permuted proteins) by the selected particles can be expressed and evaluated.

## Figure 2. Permuteins presented in the scanning permutagenesis library

Human G-CSF (ser17) protein is depicted as a string of circles, each circle corresponding to a single amino acid residue. Amino and carboxy ends of the protein are indicated. The amino acids of helical regions are indicated by medium gray balls, while the amino acids of inter-helical loops are indicated in light gray balls (See Hill et al., *Proc. Natl. Acad. Sci. USA* 90: 5167-5171, 1993). Amino ends of the permuteins made for presentation in the library are indicated in dark gray. Asterisks indicate the breakpoints of the presented permuteins which were isolated by affinity screening with cells expressing hG-CSF receptor as illustrated in 1B.

# Figure 3. Bioactivity of permuteins identified by affinity screening of the scanning permutagenesis library

Individual permuteins were expressed transiently in mammalian cells. Permeation molecules in the culture supernatants were quantitated by ELISA, and the proliferative activity of clones was determined using BAF-3-cells dependent on G-CSF for growth. The horizontal axis indicate concentration of protein and the vertical axis indicate incorporation of tritiated thymidine.

#### **Definitions**

The following is a list of abbreviations and the corresponding meanings as used interchangeably herein:

g = gram(s)

mg = milligram(s)

ml and mL = milliliter(s)

RT = room temperature

ug and µg = microgram(s)

uL and µl = microliter(s)

The following is a list of definitions of various terms used herein:

The term "permutein" means a circularly-permuted protein: a protein in which the amino and carboxy ends of the parent protein are joined together by a peptide linker sequence of zero or more amino acids. The amino and carboxy ends of the permuted protein occur at amino acids within the parental sequence.

5

The terms "chemical ligation" and "conjugation" mean a chemical reaction which covalently links two similar or dissimilar functional groups together intramolecularly or intermolecularly.

10

The term "peptide linker' means a compound which forms a carboxamide bond between two groups having one or more peptide linkages (CONH-) and serves as a connector for the propose of amelioration of the distance or space orientation between two molecules.

The term "native sequence" refers to an amino acid or nucleic acid sequence which is identical to a wild-type or native form of a gene or protein.

15

20

25

30

The terms "mutant amino acid sequence," "mutant protein", "variant protein", "mutein", or "mutant polypeptide" refer to a polypeptide having an amino acid sequence which varies from a native sequence due to amino acid additions, deletions, substitutions, or all three, or is encoded by a nucleotide sequence from an intentionally-made variant derived from a native sequence.

## Detailed description of the invention

The present invention encompasses circularly permuted-proteins of the

## Determination of the amino and carboxyl termini of permuteins

formula C'-L'-N' prepared by phage display techniques. The polypeptide can be joined either directly or through a linker segment. The term "directly" defines permuteins in which the polypeptide ends are joined without a linker. Thus L' represents a chemical bond or a linker, preferably a polypeptide segment to which both C' and N' are joined, wherein C' is comprised of a segment derived from the carboxy portion of the parent protein and N' is comprised of a segment derived from the amino terminal portion of a parent protein represented by the general formula N'-C'. Preferably, N' and C' in the permuted protein C'-L'-N' are the same length as in the parent protein N'-C', but each may be independently shorter or longer depending on the desired structural characteristics of the permutein.

Most commonly L' is a linear peptide in which C' and N' are joined by amide

10

15

20

25

30

bonds, linking  $C^i$  to the amino terminus of  $L^i$  and carboxy terminus of  $L^i$  to the amino terminus of  $N^i$ .

Additional peptide sequences may also be added to facilitate purification or identification of permuteins (e.g., poly-His). A highly antigenic peptide may also be added that would enable rapid assay and facile purification of the permuteins by a specific monoclonal antibody.

#### Determination of the linker

The linking group (L¹) is generally a polypeptide of between 1 and 500 amino acids in length. The linkers joining the two molecules are preferably designed to (1) allow the two molecules to fold and act independently of each other, (2) not have a propensity for developing an ordered secondary structure which could interfere with the functional domains of the two proteins, (3) have minimal hydrophobic characteristics which could interact with the functional protein domains and (4) provide steric separation of C¹ and N¹ such that C¹ and N¹ could interact simultaneously with their corresponding receptors on a single cell. Typically surface amino acids in flexible protein regions include Gly, Asn and Ser. Virtually any permutation of amino acid sequences containing Gly, Asn and Ser would be expected to satisfy the above criteria for a linker sequence. Other neutral amino acids, such as Thr and Ala, may also be used in the linker sequence. Additional amino acids may also be included in the linkers due to the addition of unique restriction sites in the linker sequence to facilitate construction of the multi-functional proteins.

Preferred L<sup>1</sup> linkers of the present invention include sequences selected from the group of formulas:

(SEQ ID NO: 1) through SEQ ID NO: 268)

Other linkers are also contemplated by the invention. The present invention is, however, not limited by the form, size or number of linker sequences employed. The only requirement of the linker is that it does not functionally interfere with the folding and function of the individual molecules of the multifunctional protein.

#### Utility of permuteins

Permuteins of the present invention may exhibit useful properties such as having similar or greater biological activity when compared to a single factor or by having improved half-life or decreased adverse side effects, or a combination of these properties.

Permuteins which have little or no activity maybe useful as antigens for the production of antibodies for use in immunology or immunotherapy, as probes or as intermediates used to construct other useful permuteins.

The permuteins of the present invention may have an improved therapeutic profile as compared to their parent molecules. For example, some permuteins of the present invention may have a similar or more potent activities relative to other compounds or proteins without having a similar or corresponding increase in side-effects. This is particularly true of multifunctional or fusion protein therapeutics, where permutation may relieve steric and other hindrances that impair the activity of the parent fusion molecules (see Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* 55:3357-3363, 1995, for examples).

15

5

10

A general utility of permuteins is in the area of nanoscale devices described alternatively as "nanobiological" or "nanobiotechnological." These are nanoscale devices containing both precise structure nanomaterials and biological functional components (such as proteins). Nanodevices have been the subject of several reviews (Lee, S.C., Trends in Biotechnology, 16: 239-240, 1998).

20

25

Nanobiological/nanobiotechnological devices generally contain proteins covalently coupled to polymers or other non-biological precise structure materials. Issues of steric and other interferences with protein activity are applicable to proteins in nanobiological/nanobiotechnological devices and are highly analogous to the issues with multifunctional/fusion proteins discussed above. Protein permutation is fully expected to offer a viable approach to deal with these considerations, just as it does in the case of fusion proteins (Kreitman, R. J. et al., *Proc Natl Acad Sci USA* 91: 6889-6893, 1994; Kreitman, R. J. et al., *Cancer Res.* 55:3357-3363, 1995).

#### Examples

30

The following examples will illustrate the invention in greater detail, although it will be understood that the invention is not limited to these specific examples. Various other examples will be apparent to the person skilled in the art after reading the present disclosure without departing from the spirit and scope of

the invention. It is intended that all such other examples be included within the scope of the appended claims.

#### General Materials and Methods

General methods of cloning, expressing, and characterizing proteins are found in T. Maniatis et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982, and references cited therein, incorporated herein by reference; and in J. Sambrook et al., *Molecular Cloning, A Laboratory Manual*, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory, 1989, and references cited therein, incorporated herein by reference.

10

5

Unless noted otherwise, all specialty chemicals were obtained from Sigma Chemical Company (St. Louis, MO). Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs (Beverly, MA) or Boehringer Mannheim (Indianapolis, IN).

#### Strains, plasmids, and bacteriophage

15

The bacterial strains used in these studies are listed in Table 1. Plasmids and bacteriophage used or constructed in this study are listed in Tables 2 and 3, respectively.

20

Phage and phagemid stocks were made and manipulated as described (Kay, B.K., Winter, J., and McCafferty, J., Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996; Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997).

#### Transformation of E. coli strains

E. coli strains (Table 1), such as DH5α<sup>TM</sup> (Life Technologies, Gaithersburg, MD) and TGI (Amersham Corp., Arlington Heights, IL) are used for transformation of ligation reactions and are the hosts used to prepare plasmid DNA for transfecting mammalian cells. E. coli strains, such as JM101 (Yanisch-Perron et al., Gene, 33: 103-119, 1985) and MON105 (Obukowicz et al., Appl. and Envir. Micr., 58: 1511-1523, 1992) can be used for expressing the multi-functional proteins of the present invention in the cytoplasm or periplasmic space.

30

OCHROCEAN IA-

PHEDOCID: AND

25

DH5 $\alpha^{\text{TM}}$  Subcloning efficiency cells are purchased as competent cells and are ready for transformation using the manufacturer's protocol, while both  $E.\ coli$  strains TG1 and MON105 are rendered competent to take up DNA using a CaCl,

10

15

20

25

30

35

method. Typically, 20 to 50 mL of cells are grown in LB medium (1% Bactotryptone, 0.5% Bacto-yeast extract, 150 mM NaCl) to a density of approximately 1.0 optical density unit at 600 nanometers (OD600) as measured by a Baush & Lomb Spectronic spectrophotometer (Rochester, NY). The cells are collected by centrifugation and resuspended in one-fifth culture volume of CaCl<sub>2</sub> solution (50 mM CaCl<sub>2</sub>, 10 mM Tris-Cl, pH7.4) and are held at 4°C for 30 minutes. The cells are again collected by centrifugation and resuspended in one-tenth culture volume of CaCl<sub>2</sub> solution. Ligated DNA is added to 0.2 mL of these cells, and the samples are held at 4°C for 30-60 minutes. The samples are shifted to 42°C for two minutes and 1.0 mL of LB is added prior to shaking the samples at 37°C for one hour. Cells from these samples are spread on plates (LB medium plus 1.5% Bacto-agar) containing either ampicillin (100 micrograms/mL, ug/mL) when selecting for ampicillin-resistant transformants, or spectinomycin (75 ug/mL) when selecting for spectinomycin-resistant transformants. The plates are incubated overnight at 37°C.

Colonies are picked and inoculated into LB plus appropriate antibiotic (100 ug/mL ampicillin or 75 ug/mL spectinomycin) and are grown at 37°C while shaking.

## DNA isolation and characterization

DNA constructs were made and propagated in *E. coli* using standard molecular biology techniques (Sambrook, J. et al., *Molecular Cloning*, *A Laboratory Manual*, 2<sup>ed</sup> edition, Cold Spring Harbor Laboratory, 1989).

Plasmid DNA can be isolated by a number of different methods and using commercially available kits known to those skilled in the art. Plasmid DNA is isolated using the Promega Wizard<sup>TM</sup> Miniprep kit (Madison, WI), the Qiagen QIAwell Plasmid isolation kits (Chatsworth, CA) or Qiagen Plasmid Midi or Mini kit. These kits follow the same general procedure for plasmid DNA isolation. Briefly, cells are pelleted by centrifugation (5000 x g), the plasmid DNA released with sequential NaOH/acid treatment, and cellular debris is removed by centrifugation (10000 x g). The supernatant (containing the plasmid DNA) is loaded onto a column containing a DNA-binding resin, the column is washed, and plasmid DNA eluted. After screening for the colonies with the plasmid of interest, the E. coli cells are inoculated into 50-100 ml of LB plus appropriate antibiotic for overnight growth at 37°C in an air incubator while shaking. The purified plasmid DNA is used for DNA sequencing, further restriction enzyme digestion, additional

10

15

20

25

30

subcloning of DNA fragments and transfection into  $E.\ coli$ , mammalian cells, or other cell types.

#### Sequence confirmation

DNA sequence analysis was performed using the Genesis 2000 DNA analysis system using standard methods (Prober et al., Science 238: 336-341, 1987).

Purified plasmid DNA is resuspended in dH2O and its concentration is determined by measuring the absorbance at 260/280 nm in a Bausch and Lomb Spectronic 601 UV spectrometer. DNA samples are sequenced using ABI PRISM™ DyeDeoxy™ terminator sequencing chemistry (Applied Biosystems Division of Perkin Elmer Corporation, Lincoln City, CA) kits (Part Number 401388 or 402078) according to the manufacturer's suggested protocol usually modified by the addition of 5% DMSO to the sequencing mixture. Sequencing reactions are performed in a DNA thermal cycler (Perkin Elmer Corporation, Norwalk, CT) following the recommended amplification conditions. Samples are purified to remove excess dye terminators with Centri-Sep™ spin columns (Princeton Separations, Adelphia, NJ) and lyophilized. Fluorescent dye labeled sequencing reactions are resuspended in deionized formamide, and sequenced on denaturing 4.75% polyacrylamide-8M urea gels using ABI Model 373A and Model 377 automated DNA sequencers. Overlapping DNA sequence fragments are analyzed and assembled into master DNA contigs using Sequencher DNA analysis software (Gene Codes Corporation, Ann Arbor, MI).

## Expression of permuted proteins in mammalian cells

To obtain sufficient protein for analysis of the bioactivity of individual MPO molecules containing cphG-CSFs, DNA segments containing individual affinity-selected MPO: cphGCSFs were subcloned into a mammalian expression vector, and expressed transiently in BHK cells as described below.

The BHK-21 cell line can be obtained from the ATCC (Rockville, MD). The cells are cultured in Dulbecco's modified Eagle media (DMEM/high-glucose), supplemented to 2 mM (mM) L-glutamine and 10% fetal bovine serum (FBS). This formulation is designated BHK growth media. Selective media is BHK growth media supplemented with 453 units/mL hygromycin B (CalBiochem, San Diego, CA). The BHK-21 cell line was previously stably transfected with the HSV transactivating protein VP16, which transactivates the IE110 promoter found on

10

15

20

the plasmid pMON3359 and pMON3633 and the IE175 promoter found in the plasmid pMON3360B (Hippenmeyer, P.J. and Pegg, L.E., Curr. Opin. Biotechnol. 6: 548-552, 1995). The VP16 protein drives expression of genes inserted behind the IE110 or IE175 promoter. BHK-21 cells expressing the transactivating protein VP16 are designated BHK-VP16. The plasmid pMON1118 expresses the hygromycin resistance gene from the SV40 promoter (Highkin et al., Poultry Sci., 70: 970-981, 1991). A similar plasmid, pSV2-hph, is available from ATCC.

BHK-VP16 cells are seeded into a 60 millimeter (mm) tissue culture dish at 3 x 10° cells per dish 24 hours prior to transfection. Cells are transfected for 16 hours in 3 mL of "OPTIMEM"TM (Gibco-BRL, Gaithersburg, MD) containing 10 ug of plasmid DNA containing the gene of interest, 3 ug hygromycin resistance plasmid, pMON1118, and 80 ug of Gibco-BRL "LIPOFECTAMINE" per dish. The media is subsequently aspirated and replaced with 3 mL of growth media. At 48 hours post-transfection, media from each dish is collected and assayed for activity (transient conditioned media). The cells are removed from the dish by trypsin-EDTA, diluted 1:10, and transferred to 100 mm tissue culture dishes containing 10 mL of selective media. After approximately 7 days in selective media, resistant cells grow into colonies several millimeters in diameter. The colonies are removed from the dish with filter paper (cut to approximately the same size as the colonies and soaked in trypsin/EDTA) and transferred to individual wells of a 24 well plate containing 1 mL of selective media. After the clones are grown to confluence, the conditioned media is re-assayed, and positive clones are expanded into growth media.

#### Affinity selection and screening of phagemids

25

30

Affinity reagent used for the identification of functional MPO molecules containing cphG-CSF (MPO: cphG-CSF) species from the library were BHK cells expressing the hG-CSF receptor on their surface. The library pool was subjected to iterative affinity selection (four rounds) against BHK cells expressing the h-GCSF receptor using previously described techniques (Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997). Between rounds of selection, phage eluted from the affinity reagent were amplified in E. coli (Kay, B.K. J. Winter, and J. McCofferty, Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996).

10

15

20

25

30

#### Expression of proteins in E. coli

When large-scale quantities of recombinant protein are desirable for structure-function studies, DNA segments containing individual affinity-selected MPO:cphGCSFs are subcloned into any of a variety of bacterial plasmid expression vectors, and expressed as a cytoplasmic product or as a secreted protein in *E. coli*.

E. coli strain MON105 or JM101 harboring the plasmid of interest are grown at 37°C in M9 plus casamino acids medium with shaking in an air incubator Model G25 from New Brunswick Scientific (Edison, NJ). Growth is monitored at OD<sub>so</sub> until it reaches a value of 1.0 at which time nalidixic acid (10 mg/mL) in 0.1 N NaOH is added to a final concentration of 50 µg/mL, for cultures containing plasmids with the E. coli recA promoter driving expression of the recombinant gene. IPTG is used in place of nalidixic acid, as a chemical inducer to facilitate expression from plasmids containing the lac promoter or hybrid lac promoters. The cultures are then shaken at 37°C for three to four additional hours. A high degree of aeration is maintained throughout the culture period in order to achieve maximal production of the desired gene product. The cells are examined under a light microscope for the presence of inclusion bodies (IB). One mL aliquots of the culture are removed for analysis of protein content by boiling the pelleted cells, treating them with reducing buffer and electrophoresis via SDS-PAGE (see Maniatis et al., "Molecular Cloning: A Laboratory Manual", 1982). The culture is centrifuged (5000 x g) to pellet the cells.

#### Isolation of Inclusion Bodies

The cell pellet from a 330 mL E. coli culture is resuspended in 15 mL of sonication buffer (10 mM 2-amino-2-(hydroxymethyl) 1,3-propanediol hydrochloride (Tris-HCl), pH 8.0 + 1 mM ethylenediaminetetraacetic acid (EDTA). These resuspended cells are sonicated using the microtip probe of a Sonicator Cell Disruptor (Model W-375, Heat Systems-Ultrasonics, Inc., Farmingdale, New York). Three rounds of sonication in sonication buffer followed by centrifugation are employed to disrupt the cells and wash the inclusion bodies (IB). The first round of sonication is a 3 minute burst followed by a 1 minute burst, and the final two rounds of sonication are for 1 minute each.

#### Purification

The folded proteins can be affinity-purified using affinity reagents such as monoclonal antibodies or receptor subunits attached to a suitable matrix.

10

15

20

25

30

Purification can also be accomplished using any of a variety of chromatographic methods such as: ion exchange, gel filtration or hydrophobic chromatography or reversed phase HPLC. These and other protein purification methods are described in detail (Methods in Enzymology, Volume 182 "Guide to Protein Purification" edited by Murray Deutscher, Academic Press, San Diego, California, 1990).

## **Protein Characterization**

The purified protein is analyzed by RP-HPLC, electrospray mass spectrometry, and SDS-PAGE. The protein quantitation is done by amino acid composition, RP-HPLC, and Bradford protein determination. In some cases tryptic peptide mapping is performed in conjunction with electrospray mass spectrometry to confirm the identity of the protein.

### Baf-3/G-CSF receptor assay

Briefly, the mouse lymphoid cell line Baf3 was transfected with human granulocyte colony stimulating factor receptor (hG-CSFR) cDNA. Stable clones of Baf3 which expressed the G-CSFR and proliferated in the presence of hG-CSF were isolated and used to investigate the activity of human G-CSF receptor agonists without the influence of other human cytokine receptor responses.

The cDNA encoding hG-CSFR (a gift from Dr. Daniel C. Link (Washington University, St. Louis, MO) was released from the plasmid pEMCV. Sralpha as a HindIII/EcoRI (5' to 3') fragment, gel-purified, and inserted into the mammalian cell expression plasmid pcDNA3 (Invitrogen, San Diego, CA). This plasmid contains enhancer-promoter sequences from the immediate early gene of the human cytomegalovirus (CMV), a bovine growth hormone polyadenylation signal and transcription termination sequences, a neomycin resistance gene is present for the selection of G418 stable cell clones, and an ampicillin resistance gene for selection in E. coli. Ligation mixtures were transformed into E. coli strain TG1 [delta (lac-pro), supE, thi, hsdD5/F(traD36, proAB, lacI, lacZdeltaM15] and plasmid DNA was purified using a Qiagen Midiprep Plasmid Kit. The structure of plasmid DNAs containing hG-CSFR were confirmed by restriction enzyme analysis and by automated DNA sequence analysis using an ABI sequencing machine. One of several plasmids with the correct structure was selected and given the designation pMON30298.

Bai3 cells, maintained in complete growth medium (RPMI 1640 supplemented to 10% FBS and 10% Wehi 3B supernatant as a source for mouse IL-

10

15

20

25

30

35

3), were seeded at a subconfluent cell density of 10^5 cells/ml in growth media (RPMI 1640 5% FBS; 2 mM L-glutamine) the day prior to the electroporation. The cells were collected and rinsed twice in 10 ml serum-free RPMI 1640. The cells were diluted to 10^6/ml in serum-free RPMI and 1 ml was placed into each electroporation chamber (Gibco/BRL #1608AJ). 50 ug of plasmid DNA was added to each chamber and the chambers were incubated on ice for 30 minutes prior to electroporation. The cells were electroporated on ice at a capacitance of 800 uF, 400V, fast charge, and low ohms in a BRL CellPorator. The cells were immediately removed from the chambers and placed into 10 cm dishes containing 10 ml of growth medium. The cells were allowed to recover for 48 hr in growth media prior to selection.

After the recovery period, the cells were pelleted at 1000 rpm for 10 minutes, and resuspended into 10 ml of selection medium (growth medium containing 800 ug/ml G418 sulfate (Gibco/BRL). The cells were kept in selection media, being passaged twice weekly, until only a few viable cells could be seen in the mock transfected control cell dishes (approximately 2 weeks). After an additional 2 weeks in selection media, the cells which had been electroporated with the hG-CSFR cDNA had grown to a cell density which allowed them to be tested for proliferation in the presence of hG-CSF (Fukunaga, R. et al., EMBO J. 10 (10): 2855-2865, 1991).

The cell proliferation assay conditions are as follows: Briefly, 25,000 cells were plated in a microtiter 96 well plate with or without cytokine in IMDM medium supplemented with BSA (50 ug/ml), human transferrin (100 ug/ml), lipid (50 ug/ml) 2-mercaptoethanol (50 uM final concentration). Each well was incubated with 0.5 uCi of 'H-thymidine (16 hours) and the incorporated radioactivity was measured. Triplicate wells containing Baf3 cells were set up with 4 nM hG-CSF, 4 nM mIL-3 or media only control. Samples of different permuted proteins were tested in each assay.

## Example 1: Construction of a permutein library without a linker region

Figure 1 shows a schematic of scanning permutagenesis. A plasmid construct comprising a tandem repeat of the modified human granulocyte colony stimulating factor (hG-CSF with a serine for amino acid 17) gene joined by a sequence (GCCGG, termed a zero order linker) was generated and subcloned into the plasmid pACYC177 (Chang, A.C.Y. and S.N. Cohen, *J Bacteriol*. 134: 1141-1156, 1978) using standard molecular biology methods (Sambrook, J. et al.,

Molecular Cloning, A Laboratory Manual, 2<sup>nd</sup> edition, Cold Spring Harbor Press, New York, 1989). The resultant plasmid construct (pMON15978) was linearized by restriction digestion (SmaI) and used as a template for PCR amplification of circularly permuted hG-CSF (cphG-CSF) genes, following the method of Horlich (Horlick, R.A. et al., Protein Engineering 5: 427-431, 1992. For purposes of this demonstration of the scanning permutagenesis technique, we chose to make a limited permutein library rather than one containing every possible cphG-CSF. Figure 2 shows the position of the new amino termini for each new cphG-CSF.

Individual cphG-CSF genes were inserted into phagemid presentation

vector pCANTAB 5E (Pharmacia Biotech,) such that they were expressed as a part of a MPO species (Feng, Y., N. R. Staten, C. M. Baum, N. L. Summers, M. Caparon, S. C. Bauer, L. Zurfluh, J. P. McKearn, B. K. Klein, S. C. Lee, C. A. McWherter. 1997. Multi-functional hematopoeitic receptor agonists. World Patent Application WO 97/12985) which was in turn fused to the amino end of the phage geneIII product. The presented fusion protein contained, starting from its amino terminus, a hIL-3 receptor agonist, cphG-CSF, and the phage gene III product. The juncture

between the presented protein and the gene III product was as previously described (Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29,

1997).

20

25

5

10

15

After confirmation of the structure of each phagemid construct, phagemid particles were produced for each individual cphG-CSF-presenting species (Merlin et al., 1997). Some of these lots of particles were used to individually define the affinity properties of specific presented cphG-CSF species in analytical biopanning experiments (Caparon, M. H. et al., Molecular Diversity 1: 241-246, 1996; Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997), but all of the phage particle lots were titered and equivalent numbers of transducing units of each particle preparation were pooled together to form the scanning permutagenesis library for hG-CSF in an MPO background. Figure 2 shows the MPO: cp hG-CSF species present in the library.

30

MPO: cphG-CSF 38/37 is an example of the nomenclature used to specify the identity of individual permuted proteins. It describes a MPO molecule containing a circularly permuted human G-CSF module (with the serine 17 substitution). The first amino acid of the cphG-CSF domain is amino acid 38 of the parent protein, and the last amino acid is residue 37 of the parent.

library

MPO: cphG-CSF 38/37, is a full hG-CSF receptor agonist (McKearn, J.P., Myelorestorative activities of synthokine and myelopoietin. In Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996). It was presented on filamentous phage as a positive control to demonstrate that permuted proteins can be presented on the surface of phage particles and affinity selected. After phagemid particles were produced from this construct, they were subjected to analytical biopanning using cells expressing the hG-CSF receptor as affinity reagent.

10

5

Table 1 shows that phage presented MPO: cphG-CSF 38/37 was affinity selected by cells expressing the hG-CSF. MPO: cphGCSF 38/37-GPIII fusion was expressed, secreted and assembled into phagemid particles, and could be affinity selected by the hG-CSF receptor. Permutagenesis of a protein does not appear to impair its successful presentation.

15

Relative to typical phage display libraries, the complexities of cp libraries are low, containing perhaps hundreds to thousands of individuals. The demonstration library here contained about 50 distinct clones, as opposed to more typical phage libraries containing more than 10<sup>5</sup> individuals (reviewed in Clackson, T. and J.A. Wells, *Tibtech* 12: 173-184, 1994).

20

37 randomly chosen selectants from round 1, and fewer from subsequent rounds (17, 11 and 14 were picked from rounds 2, 3 and 4, respectively) were chosen for sequence analysis. The identity of the MPO: cphG-CSFs identified in each round is shown in Figure 2.

25

30

A total of 14 MPO: cphGCSF species were identified from the output of affinity selection (Figure 2). Most of the MPO: cphGCSF species identified from the library had new carboxy and amino termini in loop segments (9 of 14 permuteins identified), rather than in clearly defined secondary structures (See Hill et al., 1993 for the hG-CSF structure). Five selectants had termini within helical domains of hG-CSF (MPO: cphG-CSFs 13/12, 19/18, 71/70, 123/122 and 159/158). For three of these molecules (MPO: cphG-CSFs 13/12, 71/70 and 123/122) their new ends lie at the outermost ends of helices, and therefore perturbation of secondary structure caused by these permuteins may be minimal. However, MPO: cphG-CSF 19/18 and MPO: cphG-CSF 159/158 have new termini well within helix 1 and helix 4 of hG-CSF, respectively.

10

15

20

25

30

These data parallel the observations of Graf and Schachman, who developed a limited DNase I digestion method for "random" permutagenesis (Graf, R. and H. K. Schachman, *Proc Natl Acad Sci USA* 93:11591-11596, 1996). They identified two permutein species of aspartate transcarbamoylase that introduced new amino and carboxy ends into secondary structural domains and that retained biological activity. In their work, the majority of permuteins introducing ends into secondary structures (5/7 identified) were significantly diminished in activity. In contrast, we found a several permuteins that introduced helical breaks retained activity (See Below). The method used by Graf and Schechman frequently introduces point mutations, small insertions and deletions into the permuted proteins, potentially complicating the analysis of the effects of permutagenesis.

# Example 3: Biological activity of MPO: cphG-CSFs selected from the cp phage library

To obtain sufficient protein for analysis of the bioactivity of individual MPO molecules containing cphG-CSFs, DNA segments containing individual affinity-selected MPO: cphGCSFs were subcloned into a mammalian expression vector, and expressed transiently in BHK cells as described above.

The MPO: cphG-CSFs isolated from biopanning were all expressed transiently in mammalian cells and the amount of MPO: cphG-CSF in each supernatant was determined by sandwich hll-3 ELISA (Olins P.O. et al., *J. Biol. Chem.* 270: 23754-23760, 1995). The quantitated supernatants were then assayed for G-CSF receptor agonist activity in a Baf-3/G-CSF receptor assay (Figure 3, Table 4).

All but one of the transiently expressed MPO: cphG-CSF proteins exhibited G-CSF activity equivalent to or slightly better than that of the parent MPO molecule, including those MPO: cphG-CSFs with new carboxy and amino ends within helixes. The permutein encoded by pMON16021 with a breakpoint between positions 48 and 49 did not exhibit activity in the G-CSF-dependent proliferation assay. These data suggest that most of the proteins isolated from the library are competent to bind the hG-CSF receptor and produce a proliferation signal.

All references, patents, or applications cited herein are incorporated by reference in their entirety, as if written herein.

#### References

Buchwalder, A. et al., Biochemistry 31:1621-1630, 1992.

Caparon, M. H. et al., Molecular Diversity 1: 241-246, 1996.

Chang, A.C.Y. and S.N. Cohen, J Bacteriol. 134: 1141-1156, 1978.

5 Clackson, T. and J.A. Wells, Tibtech 12: 173-184, 1994.

Feng et al., J. Mol. Biol. 259: 524-551, 1996.

Feng, Y., N. R. Staten, C. M. Baum, N. L. Summers, M. Caparon, S. C. Bauer, L. Zurfluh, J. P. McKearn, B. K. Klein, S. C. Lee, C. A. McWherter. 1997. Multifunctional hematopoeitic receptor agonists. World Patent Application WO 97/12985.

Fukunaga, R., Ishizaka-Ikeda, E., Pan, C. X., Seto, Y., and Nagata, S. Functional Domains of the Granulocyte Colony-Stimulating Factor Receptor. *EMBO J.* 10 (10):2855-2865, 1991.

Goldenberg, D. P. and T. E. Creighton, J. Mol. Biol. 165: 407-413, 1983.

15 Graf, R. and H. K. Schachman, Proc Natl Acad Sci USA 93:11591-11596, 1996.

Hahn, M. et al., Proc Natl Acad Sci USA 91: 10417-10421, 1994.

Highkin et al., Poultry Sci., 70: 970-981, 1991.

Hill et al., Proc. Natl. Acad. Sci. USA 90: 5167-5171, 1993.

Hippenmeyer, P.J. and L.E. Pegg, Curr. Opin. Biotechnol. 6: 548-552, 1995.

Horlick, R.A. et al., Protein Engineering 5: 427-431, 1992.

Jelinski, L.W. Biologically related aspects of nanoparticles, nanostructured materials and nanodevices. *In* "WTEC workshop on Global Assessment of R &D Status and Trends in Nanoparticles, Nanostructured Materials and Nanodevices", International Technology Research Institute, Loyola College, Baltimore, MD., S.

25 C., 1998.

Johnson, J. and F. M. Raushel, Biochemistry 35: 10223-10233, 1996.

Kay, B.K. J. Winter, and J. McCofferty, Phage Display of Peptides and Proteins, Academic Press, San Diego, California, 1996.

Koebnik, R. and L. Kramer, J. Mol. Biol. 250: 617-626, 1995.

Kreitman, R. J. et al. Cytokine 7(4): 311-318, 1995.

5 Kreitman, R. J. et al., Proc Natl Acad Sci USA 91: 6889-6893, 1994.

Kreitman, R. J. et al., Cancer Res. 55:3357-3363, 1995.

Lee, S.C. Biotechnology for Nanotechnology. *Trends in Biotechnology*, **16**: 239-240, 1998.

Lee, S.C., Phage presentation for cytokine engineering. IBC's Second International Conference on Display Technologies, 1997.

Lin, X. et al., Protein Science 4: 159-166, 1995.

Lowman and Wells, J. Mol. Biol. 234: 564-578, 1993.

Luger et al., Protein Engineering 3: 249-258, 1990.

Luger et al., Science 243: 206-210, 1989.

Maniatis, T. et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, 1982.

McKearn, J.P. Myelorestorative activities of synthokine and myelopoietin. In Proceedings of the 1996 IBC Conference on Therapeutic Applications of Cytokines, pp. 4.3.1-4.3.18, 1996.

20 Merlin, S. et al., Applied Biochemistry and Biotechnology 67: 15-29, 1997.

Mullins, L. S. et al., J Am. Chem Soc. 116: 5529-5533, 1994.

Murray Deutscher (ed), Methods in Enzymology, Volume 182 "Guide to Protein Purification," Academic Press, San Diego, California, 1990.

Obukowicz, M. et al., Appl. and Envir. Micr., 58: 1511-1523, 1992.

25 Olins P.O. et al., J. Biol. Chem. 270: 23754-23760, 1995.

Prober, J.M. et al., Science 238: 336-341, 1987.

Protosova, N. Y. et al., Protein Engineering 7: 1373-1377, 1994.

Puri, R. K. et al., Cellular Immunol. 171: 80-86, 1996.

Rose, K. et al., Molecular Immunology 32: 1031-1037, 1995.

Sambrook, J. et al., Molecular Cloning, A Laboratory Manual, 2<sup>nd</sup> edition, Cold Spring Harbor Press, New York, 1989.

Shortle, D. and J. Sondek, Curr. Opin. Biotechnol. 6: 299-305.

Smith, G. P., Curr. Opin. in Biotechnol. 2: 668-673, 1991.

Smith, G. P., Science 228: 1315-1317, 1985.

Sondek, J. and D. Shortle, Proteins 7: 387-393, 1990.

10 Thomas, J. W. et al., Proc Natl Acad Sci USA 92: 3779-3783, 1995.

Winter, G., Drug Development Res. 33: 71-89, 1994.

Yang, Y. R. and H. K. Schachman, Proc Natl Acad Sci USA 90: 11980-11984, 1993.

Yanisch-Perron et al., Gene, 33: 103-119, 1985.

Zhang, T. et al., Biochemistry 32: 12311-12318, 1993.

### Tables

Table 1: Circularly permuted proteins

Protein	Reference
Enzymes	
T4 lysozyme	Zhang et al., Biochemistry 32:12311-12318 (1993);
	Zhang et al., Nature Struct. Biol. 1:434-438 (1995)
dihydrofolate reductase	Buchwalder et al., Biochemistry 31:1621-1630
•	(1994);
	Protasova et al., Prot. Eng. 7:1373-1377 (1995)
ribonuclease T1	Mullins et al., J. Am. Chem. Soc. 116:5529-5533
	(1994);
	Garrett et al., Protein Science 5:204-211 (1996)
Bacillus β-glucanase	Hahn et al., <i>Proc. Natl. Acad. Sci. U.S.A.</i> 91:10417 10421 (1994)
aspartate transcarbamoylase	Yang and Schachman, Proc. Natl. Acad.
	Transcarbamoylase Sci. U.S.A. 90:11980-11984 (1993)
phosphoribosyl anthranilate	Luger et al., Science 243:206-210 (1989);
isomerase	Luger et al., Prot. Eng. 3:249-258 (1990)
pepsin/pepsinogen	Lin et al., Protein Science 4:159-166 (1995)
glyceraldehyde-3-phosphate	Vignais et al., Protein Science 4:994-1000 (1995)
dehydrogenase	
ornithine decarboxylase	Li & Coffino, Mol. Cell. Biol. 13:2377-2383 (1993)
yeast phosphoglycerate	Ritco-Vonsovici et al., Biochemistry 34:16543-
dehydrogenase	16551 (1995)
Enzyme Inhibitor	
basic pancreatic trypsin inhibitor	Goldenberg & Creighton, J. Mol. Biol. 165:407-41 (1983)
Cytokines	
interleukin-1β	Horlick et al., Protein Eng. 5:427-431 (1992)
interleukin-4	Kreitman et al., Cytokine 7:311-318 (1995)

# Tyrosine Kinase Recognition Domain

 $\alpha$ -spectrin SH3 domain

Viguera et al., J. Mol. Biol. 247:670-681 (1995)

Transmembrane Protein

omp A

Koebnik & Krämer, J. Mol. Biol. 250:617-626

(1995)

Chimeric Protein

interleukin-4-Pseudomonas

exotoxin fusion molecule

Kreitman et al., Proc. Natl. Acad. Sci. U.S.A.

91:6889-6893 (1994)

Table 2: Strains

Designation	Description or Genotype	Reference/Source
DH5α <sup>TM</sup>	F, phi80 dlacZdeltaM15, delta(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17 (rk,mk*), phoA, supE44, lambda-, thi-1, gyrA96, relA1	Life Technologies, Rockville, Maryland
JM101 (ATCC# 33876)	delta (pro lac), supE, thi, F'(traD36, proA*B*, lacI*, lacZdeltaM15)	Yanisch-Perron et al., <i>Gene</i> , 33: 103-119, 1985
MON105 (ATCC# 55204)	F, lambda-,IN (rrnD, rrnE)1, rpoD <sup>+</sup> , rpoH358	Obukowicz et al., <i>Appl. and Envir. Micr.</i> , 58: 1511-1523, 1992
MON208	W3110 rpoH358, lacI <sup>q</sup> , ompT::kan	Alan Easton
TG1	delta(lac-pro), supE, thi-1, hsdD5/F"(traD36, proA*B*, lacIq, lacZdeltaM15)	Amersham Corp., Arlington Heights, Illinois
W3110	IN (rrnD-rrnE)1, rph1	Lab collection

Table 3: Plasmids

Plasmid	SEQ ID NO.	Selectable Marker	Description	Source
pACYC177		Kan <sup>R</sup> Amp <sup>R</sup>	Plasmid with multiple cloning sites and two selectable markers	Chang, A.C.Y. and S.N. Cohen, J Bacteriol. 134: 1141- 1156, 1978
pMON15978		AmpR	Plasmid construct comprising a tandem repeat of the modified human granulocyte colony stimulating factor (hG-CSF with a serine for amino acid 17) gene joined by a sequence (GCCGG, termed a zero order linker), subcloned into the plasmid pACYC177	This work
pCANTAB 5E		Amp <sup>R</sup>	Phage display vector containing lac promoter operably linked to fd gene 3 signal sequence, a linker region, an E tag, and an fd gene 3 structural gene all cloned into the vector backbone of pUC119 containing ColE1 ori, the beta lactamase resistance gene, and an M13 ori.	Pharmacia Biotech, Piscataway, NJ
pMON16016		AmpR	Phagemid presentation vector pCANTAB 5E derivation containing inserted individual cphG-CSF gene such that it was expressed as a part of an	This work

MPO species, fused in turn to the amino terminus end of the phage geneIII product. The first amino acid of the cphG-CSF domain is amino acid 1 of the parent, and the last amino acid is residue 174 of the parent. The zero order linker is attached at the carboxyl end of amino acid 174.

the parent.

pMON16017

 $Amp^R$ 

Identical to pMON16016 except This work the first amino acid of the cphG-CSF domain is amino acid 3 of the parent, and the last amino acid is residue 2 of

pMON16029

 $Amp^R$ 

Identical to pMON16016 except This work the first amino acid of the cphG-CSF domain is amino acid 7 of the parent, and the last amino acid is residue 6 of the parent.

pMON16030

 $Amp^R$ 

Identical to pMON16016 except This work the first amino acid of the cphG-CSF domain is amino acid 9 of the parent, and the last amino acid is residue 8 of the parent.

pMON16018

 $Amp^R$ 

Identical to pMON16016 except This work the first amino acid of the cphG-CSF domain is amino acid 11 of the parent, and the last amino acid is residue 10 of the parent.

pMON16019	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 13 of the parent, and the last amino acid is residue 12 of the parent.	This work
pMON16031	Amp <sup>R</sup>	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 15 of the parent, and the last amino acid is residue 14 of the parent.	This work
pMON16020	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 19 of the parent, and the last amino acid is residue 18 of the parent.	This work
pMON16032	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 22 of the parent, and the last amino acid is residue 21 of the parent.	This work
pMON16033	$^{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 27 of the parent, and the last amino acid is residue 26 of the parent.	This work
pMON16034	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 31 of the parent, and the	This work

last amino acid is residue 30 of the parent.

Identical to pMON16016 except This work  $Amp^R$ pMON16035 the first amino acid of the cphG-CSF domain is amino acid 35 of the parent, and the last amino acid is residue 34 of the parent. Identical to pMON16016 except This work  $_{\mathbf{Amp}}^{\mathbf{R}}$ **PMON16036** the first amino acid of the cphG-CSF domain is amino acid 37 of the parent, and the last amino acid is residue 36 of the parent. Identical to pMON16016 except This work  $Amp^R$ pMON16037 the first amino acid of the cphG-CSF domain is amino acid 38 of the parent, and the last amino acid is residue 37 of the parent. Identical to pMON16016 except This work pMON16038  $Amp^R$ the first amino acid of the cphG-CSF domain is amino acid 39 of the parent, and the last amino acid is residue 38 of the parent. Identical to pMON16016 except  $\mathbf{Amp}^{\mathbf{R}}$ pMON16039 the first amino acid of the cphG-CSF domain is amino acid 43 of the parent, and the last amino acid is residue 42 of the parent. Identical to pMON16016 except This work Amp<sup>R</sup> pMON16040 the first amino acid of the

cphG-CSF domain is amino acid 45 of the parent, and the last amino acid is residue 44 of the parent.

last amino acid is residue 55 of

the parent.

Identical to pMON16016 except This work  $\mathbf{Amp}^{\mathbf{R}}$ pMON16041 the first amino acid of the cphG-CSF domain is amino acid 47 of the parent, and the last amino acid is residue 46 of the parent. Identical to pMON16016 except  $Amp^R$ pMON16022 the first amino acid of the cphG-CSF domain is amino acid 49 of the parent, and the last amino acid is residue 48 of the parent. Identical to pMON16016 except This work  $Amp^R$ pMON16042 the first amino acid of the cphG-CSF domain is amino acid 51 of the parent, and the last amino acid is residue 50 of the parent. Identical to pMON16016 except This work  $Amp^R$ pMON16043 the first amino acid of the cphG-CSF domain is amino acid 53 of the parent, and the last amino acid is residue 52 of the parent. Identical to pMON16016 except This work  $Amp^R$ pMON16044 the first amino acid of the cphG-CSF domain is amino acid 56 of the parent, and the

pMON16023	<sub>Amp</sub> R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 60 of the parent, and the last amino acid is residue 59 of the parent.	This work
pMON16045	Amp <sup>R</sup>	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 64 of the parent, and the last amino acid is residue 63 of the parent.	This work
pMON16024	$\mathbf{Amp}^{\mathbf{R}}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 67 of the parent, and the last amino acid is residue 66 of the parent.	This work
pMON16046	$_{\mathbf{Amp}}^{\mathbf{R}}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 69 of the parent, and the last amino acid is residue 68 of the parent.	This work
pMON16025	$_{\mathbf{Amp}}\mathbf{R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 71 of the parent, and the last amino acid is residue 70 of the parent.	This work
pMON16047	Amp <sup>R</sup>	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 73 of the parent, and the	This work

		last amino acid is residue 72 of the parent.	
pMON16048	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 84 of the parent, and the last amino acid is residue 83 of the parent.	This work
pMON16049	$_{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 98 of the parent, and the last amino acid is residue 97 of the parent.	This work
pMON16050	$_{f Amp}^{f R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 100 of the parent, and the last amino acid is residue 99 of the parent.	This work
pMON16051	<sub>Amp</sub> R	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 102 of the parent, and the last amino acid is residue 101 of the parent.	This work
pMON16052	$_{ m Amp}^{ m R}$	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 112 of the parent, and the last amino acid is residue 111 of the parent.	This work
pMON16053	$_{\mathrm{Amp}}^{\mathrm{R}}$	Identical to pMON16016 except	This work

the first amino acid of the

DEIGDOOID, 4840 004000EA4 1A-

cphG-CSF domain is amino acid 121 of the parent, and the last amino acid is residue 120 of the parent.

of the parent. Identical to pMON16016 except This work pMON16026  $Amp^R$ the first amino acid of the cphG-CSF domain is amino acid 123 of the parent, and the last amino acid is residue 122 of the parent. Identical to pMON16016 except pMON16027 This work  $Amp^R$ the first amino acid of the cphG-CSF domain is amino acid 125 of the parent, and the last amino acid is residue 124 of the parent. Identical to pMON16016 except pMON16054  $_{\rm Amp}^{\rm R}$ the first amino acid of the cphG-CSF domain is amino acid 133 of the parent, and the last amino acid is residue 132 of the parent.

pMON16055

AmpR

Identical to pMON16016 except This work the first amino acid of the cphG-CSF domain is amino acid 142 of the parent, and the last amino acid is residue 141

of the parent.

of the parent.

pMON16056

AmpR

Identical to pMON16016 except

the first amino acid of the

cphG-CSF domain is amino

acid 143 of the parent, and the

last amino acid is residue 142

pMON16057	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 147 of the parent, and the last amino acid is residue 146 of the parent.	This work
pMON16028	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 159 of the parent, and the last amino acid is residue 158 of the parent.	This work
pMON16058	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 168 of the parent, and the last amino acid is residue 167 of the parent.	This work
pMON16059	AmpR	Identical to pMON16016 except the first amino acid of the cphG-CSF domain is amino acid 170 of the parent, and the last amino acid is residue 169 of the parent.	This work

10

Table 4: Analytical biopanning

Before receptor*	After receptor*	Enrichment		
1/6.6x10	1/6.5x10 <sup>-1</sup>	990-fold		

<sup>\*</sup> Amp<sup>R</sup>/Kan<sup>R</sup> resistant colonies

Analytical biopanning shows that MPO molecules containing permuted hG-CSF domains can be presented and affinity selected in a hG-CSF receptor dependent fashion. A mixture of phagemids presenting MPO: cphG-CSF 38/37 (ampicillin resistant) and M13k07 (kanamycin resistant) were exposed to BHK cells with or without the hG-CSF receptor on their surface, washed and eluted from the cell surface. Eluted phage were introduced into *E. coli* and the transfected cells were plated on media containing kanamycin or ampicillin. The ratio of ampicillin resistant to kanamycin resistant particles were determined prior to and following exposure to receptor by counting resistant colonies.

Table 5: Activity of selected permuteins

Plasmid	Permutein breakpoint in G-CSF amino acid sequence	Activity in G-CSF- dependent proliferation assay
pMON16017	3/2	+
pMON16018	11/10	· •
pMON16019	13/12	+
pMON16020	19/18	<b>+</b>
pMON16021	49/48	<del>-</del> - 0
pMON16022	60/59	+
pMON16023	67/66	+
pMON16024	69/68	+
pMON16025	71/70	z <b>+</b>
pMON16026	123/122	+
pMON16027	125/124	<b>.</b>
pMON16028	159/158	+

Table 6: SEQ ID Number/SEQ ID Name Correlation

EEQ ID	SEQ ID	Sequence									•
10.							776	GGC	CCT	GCC	AGC
1.	FGS1	cececec	ACATG ACATG	TCT	ACA CCA	TTG TTG	GGC	CCI	ecc.	AGC	TCC
2.	FGS2 FGS3	CCCCCCC	ACATG	TCT	110	GGC	CCT	GCC	AGC	1CC	CTG
3. 4.	PGS4	CCCCCCC	ACATG	TCT	GGC	CCI	CCC	AGC	TCC	CIG	ccc
5.	FGS5	CCCCCCC	ACATG	TCT	CCI	CCC	AGC TCC	CIG	CCC	CAG	CAG AGC
6.	FGS6	CCCCCCC.	ACATG	TCT	GCC AGC	JCC VCC	C10	-000	CAG	AGC	TTC
7.	FGS7	CCCCCCC	ACATG ACATG	TCT	TCC	CTG	222	CAG .	AGC	110	CTG
8. 9.	FGS8 FGS9	CCCCCCC	ACATG	TCT	CTG	CCC .	CAG	AGC	isc	CIC	CTC
10.	PGS10	CGCGCGC	ACATG	TCT	CCC	CAG	AGC	TTC	. CTG	CTC	AAG TCT
11.	PGS11	CCCCCCC	ACATG	TCT	CAG	AGC .	TTC	CIC.	-CTC	TCT	TTA
12.	FGS12	CCCCCCC	ACATG ACATG	aca aca	AGC TTC	-CIG	CIC	AAG	TCT	TTA	CAG
13. 14.	PGS13 PGS14	CCCCCCC	ACATG	TCT	CTG	CTC	AAG	TCT	TTA	GAG	CAA
15.	PGS15	CCCCCCC	ACATG	TCT	CTC	AAG	TCT	TTA	GAG	CXX	GTG
16.	FGS16	CCCCCCC	ACATG	TCT	AAG	TCT	TTA GAG	GAG CAA	CAA GTG	GTG AGG	AGG AAG
17.	PGS17	CCCCCCC	ACATG	TCT	TCT TTA	TTA GAG	CAA	GTG	AGG	AAG -	ATC
18.	FGS18 FGS19	CCCCCCC	ACATG ACATG	TCT	GAG	CAA	CTG	AGG	AAG	ATC	CAG
19. 20.	PG520	CCCCCCC	ACATG	TCT	CAA	GTG	AGG	AAG	ATC	CAG	GGC
21.	FGS21	CCCCCCC	ACATG	TCT	CIC	AGG	AAG	ATC	.CAG GGC	GGC	GAT
22.	FGS22	cececec	ACATG	1CI	AGG AAG	AAG ATC	ATC CAG	CAG	GAT	GGC	GCA
23.	FG523	CCCCCCC	ACATG ACATG	TCT	ATC	CAG	GGC	GAT	GGC	GCA	GCG
24. 25.	PGS24 PGS25	COCCCCC	ACATG	TCT	CAG	GGC	GAT	GGC	GCA	·ece	CTC
26.	FGS26	CGCGCGC	ACATG	TCT	GGC	GAT	GGC	GCX	GCG	CAG	CAG
27.	PGS27	CCCCCCC	ACATG	101	GAT	GGC GCA	GCA	C.L.C.	CTC	GAG	YYC .
28.	FGS28	COCCCCC	ACATG ACATG	TCT	GCA	GCG	CTC	CAG	GAG	AAG	CTG
, 29. 30.	FGS29 FGS30	COCOCOC	ACATG	TCT	GCG	CTC	CAG	GAG	AAG	CTG	TCI
31.	PGS31	CCCCCCC	ACATG	TCT	CIC	CAG	GAC	AAG	·CTG	TGT	200
32.	PGS32	CCCCCCC	ACATG	TCT	CAG GAG	GAG .AAG	AAG CTG	aca.	GCC	ACC	ACC TAC
33.	FGS33	CCCCCCCC.	acatg acatg	TCT	AAG	·CTG	161	ecc.	ACC	TAC	AAG
34.	FG534 FG535	CCCCCCC	ACATG	TCT	CTG	TGT	GCC	ACC	TAC	AAG	CTG
35. 36.	FGS36	CCCCCC	ACATG	TCT	TGT	GCC	ACC	TAC	AAG	CIG	CAC
37.	FGS37	CCCCCCC	ACATG	TCT	GCC	ACC	TAC AAG	AAG CTG	TGC	CAC	CCC
38.	PGS38	CCCCCCC	ACATG ACATG	TCT	ACC TAC	AAG	CTG	1GC	CAC	~000	GAG
39. 40.	PGS39 PGS40	CCCCCCC	ACATG	TCT	AAG	CTG	TGC	CAC.	ccc	GAG	GAG
41.	PGS41	CCCCCCC	ACATG	TCT	CIG	TGC	CXC	CCC	GAG	GAG	CZC
42.	FGS42	CCCCCCC	ACATG	TCT	TGC	CAC	ccc	GAG	GAG CTG	CTG	CLC
.43.	FGS43	CCCCCCC	ACATG	TCT	CAC	CCC	GAG	-C16	GTG	CIG	crc
44.	PGS44 PGS45	CCCCCCC	ACATG ACATG	TCT	GAG	GAG	CIG	GTG	CIG	CIC	GGA
45. 46.	FGS45	CCCCCCC	ACATG	TCT	GAG	CIG	CLC .	CTG	-CTC	GGA	CXC
47.	PGS47	CCCCCCC	ACATG	ICI	CIC	GTG .	-CIG	CTC	CAC	TCT	ere ere
48.	FGS48	CCCCCCC	ACATG	TCT	CLC CLC	CIC CIG	GGA	CAC	TCT	CTG	GGC
49.	FGS49	CCCCCCC	ACATG ACATG	101	CIC	GGA	CAC	TCT	CTG	·ccc	ATC
· 50.	PGS51	CCCCCCC	ACATG	TCT	GGA	CAC	TCT	CTG	GGC	ATC	-000
52.	PGS52	CCCCCCC	ACATG	TCT	CAC	TCT	CTG	GGC	ATC -CCC	TGG	TGG GCT
53.	FGS53	CCCCCCC	ACATG	TCT	CIG	CTG	GGC ATC	ATC CCC	TGG	GCT	ccc
54.	FGS54 FGS55	CCCCCCC	ACATG ACATG	TCT	GGC	ATC	CCC	TGG	GCT	CCC	CTG
55. 56.	PUS5 5	CCCCCCC	ACATG	ICI	ATC	ccc	TGG	GCT	-ccc	CTG	AGC
57.	P3557	CCCCCCC	ACATG	TCT	ccc	TGG	GCT	ccc	CTG	AGC	TGC
58.	PGSSE	CCCCCCC	ACATG	TCT	TGG	CCC	CLC CCC	AGC	AGC TCC	TCC	CCC
59.	FGS59	COCCOCC	ACATG ACATG	TCT	CCC	CTG	AGC	TCC	TGC	ccc	AGC
60 61.	PG560 PG561	CCCCCCC	ACATG	TCT	CIG	AGC	1CC	TGC	CCC	AGC	CAG
62.		CGCGCGC	ACATG	TCT	AGC	TCC	TGC	ccc	AGC	CAG	ecc
63.		CCCCCCC	ACATG	TCT	TCC	TSC	CCC	AGC	CAG GCC	CTG CTG	CAG
64.		CCCCCCC	ACATG	303	TGC	AGC AGC	AGC CAG	GCC	CTG	CAG	CIG
65. 66.		CCCCCCC CCCCCCC	ACATG ACATG	TCT	AGC	CAG	ecc	CTG	CAG	-CTG	GCA
67.		COCOCOC	ACATG	TCT	CAG	- GCC	CIG	CAG	CLC	GCA	GGC
68.		=======================================	ACATG	TCT	GCC	CTG	CAS	CTG	GCA	GGC	760
. 69 .	FGS69	CECCECE	ACATG	TCT	CAC	CXG	CTC GCA	GCA GGC	GGC TGC	17GC	TTG AGC
70.		CCCCCCC	ACATO	TCT-	CXG	CTG GCA	GGC	TGC	776	AGC	CAL
72.	FGS7.	2525250	ACATG ACATG	TCT	GCA	GSC	TOC	176	AGC	CAA	CTC
73.		CCCCCCC	ACATG	TCT	GGC	100	TTG	AGC	CAL	CTC .	CAT
74.		cececec	ACATG	TCT	TGC		AGC	CAA	CTC	CAT	کتر محد
75.	. FGS75	cececec	ACATG -	707	7779	AGC	CYY	CTC	CAT AGC	AGC GGC	CTT
76.		cecesse	acate acate	रटा रटा	بند. دند	.TXX	CTC CAT	ASS	GSC	212	770
77. 78.		6666666	AZATG	TCT	***	CAT	AGC	GSC	^C1.1	***	CIC
75.		0000000	ACATO	727		AGC	asc	CTT	770	. 625	TAC

										•	
BC.	PGS80	CCCCCCC	ACATG	161	AGC	SGC	cir	220	222	TAC	·CAG
81.	FGS81	cececec	ACATG	TCT	GGC	CIT	1.c	CIC	TAC	CAG	·GGG
62.	FG582	CCCCCCC	ACATG	TCT	·CII	TTC	CTC	TAC	CAG	GGG	CTC
83.	FGS83	CCCCCCC	ACATG	TCT	TTC	CTC	TAC	CAG	GGG	CTC	CTG
84.	FGS84	cececec	ACATG	TCT	-C1C	TAC	CAG	CGG	CIC	CTG	CAG
85.	PGS85	CCCCCCC	ACATG	TCT	TAC	CAG	GGG	CIC	CTG	CAG	GCC
86.	PGS86	CGCGCGC	ACATG	TCT	-CAG	GGG	CTC	CTG	CAG	CCC	CTG
87.	FGS87	CCCCCCC	ACATG	TCT	GGG	CTC	CTG	CAG	GCC	CTG	GAA
		CCCCCCC	ACATG	TCT	CTC	CTG	CAG	GCC	CTG	GAA	GGG
68.	PGS88										
	PGS89	CCCCCCC	ACATG	TCT	CIG	CAG	GCC	CIG	GAA	CCC	ATA
90.	FGS90	CCCCCCC	ACATG	TCT	CAG	GCC	CTG	GAA	CCC	ATA	TCC
91.	FGS91	cececec	ACATG	TCT	GCC	CIG	GAA	CGG	ATA	TCC	ccc
92.	FG592	CCCCCCC	ACATG	TCT	CIC	GAA	CCC	ATA	TCC	CCC	GAG
93.	PGS93	CGCGCGC	ACATG	101	CYY	GGG	ATA	TEE	-ccc	GAG	TIG
94.	PGS94	CCCCCCC	ACATG	TCT	GGG	ATA	TCC	CCC	GAG	110	GGT
95.	PG595	CCCCCCC	ACATG	TCT	ATA	TCC	CCC	GAG	TTG.	CCT	222
96.	FG596	CCCCCCC	ACATG	TCT	TCC	CCC	GAG	TTG	GCT	ecc '	ACC
97.	PGS97	CCCCCCC	ACATG	TCT	CCC	GAG	776	CUT	-ccc	ACC	776
98.	PGS98	cececec	ACATG	TCT	GAG	776	GGT	ccc	ACC	77G	GAC
		CCCCCCC	ACATG	1C1	TTG	GGT	ccc	ACC	TTG	GAC	ACA
99.	PGS99						ACC				
100.	FG5100	CCCCCCC	ACATG	TCT	GGT	CCC		TTG	GAC	ACA	CTG
101.	PGS101	cececec	ACATG	TCT	ccc	ACC	TTG	GAC	ACA	CTG	-CAG
102.	PGS102	CCCCCCC	ACATG	TCT	ACC	TTG	GAC	ACA	CIG	CAG	CIG
103.	PGS103	CGCGCGC	ACATG	TCT	TTG	GAC	ACA	-Cit2	CAG	CTG	GAC
104.	PGS104	CCCCCCC	ACATG	TCT	GAC	ACA	CIG	CAG	CTG	GAC	·CIC
105.	PGS105	CGCGCGC	ACATG	TCT	ACA	CTG	CAG	Cic	GAC	CIC	GCC
106.	PGS106	CCCCCCC	ACATG	TCT	CTG	CAG	CTG	GAC	GTC	GCC	GAC
107.	PGS107	CCCCCCC	ACATG	TCT	CAG	CTG	GAC	GTC	GCC	GAC	TTT
108.	PGS108	CCCCCCC	ACATG	TCT	CTG	GAC '	GTC.	GCC	GAC	TIT	GCC
109.	FGS109	CGCGCGC	ACATG	TCT	GAC	GTC	GCC	GAC	111	GCC	ACC
110.	FGS110	CCCCCCC	ACATG	TCT	GTC .	GCC	GAC	. 277	GCC	ACC	ACC
	FGS111	CCCCCCC	ACATG	TCT	CCC	GAC	TIT .	GCC	ACC	ACC	ATC
111.							GCC				TGG
112.	FGS112	cececec	ACATG	TCT	GAC	TII		YCC .	ACC	ATC	
113.	PGS113	CCCCCCC	ACATG	TCT	III	GCC	ACC	YCC	ATC	TGG	CAG
114.	FGS114	CCCCCCC	ACATG	ICI	GCC	ACC	ACC	ATC	TGG	CAG	CAG
115.	FGS115	CCCCCCC	ACATG	TCT	ACC	ACC	ATC	200	CAG	CAG	ATG
116.	FGS116	CCCCCCC	ACATG	TCT	ACC	ATC	TOG	CAG	CAG	ATG	GAA
117.	FG5117	CCCCCCC	ACATG	TCT	ATC	TGG	CAG	CAG	ATG	GAA	GAA
118.	FGS118	CCCCCCC	ACATG	TCT	TGG	CAG	CAG	ATG	GAA	GAA	-CTG
119.	FGS119	CGCGCGC	ACATG	TCT	CAG	CXC	ATG	GAA	GAA	CTG	. GGA
120.	FGS120	CCCCCCC	ACATG	TCT	CAG	ATG	GAA	GAA	CTG	GGA	ATG
121.	FG5121	CCCCCCC	ACATG	TCT	ATG	GAA	GAA	·CZG	GGA	ATG	GCC
122.	PG5122	CCCCCCC	ACATG	TCT	GAX	GAA	CTG	GGA	ATG	GCC	CCT
123.	PG5123	CCCCCCC	ACATG	TCT	GAA	CTG	GGA	ATG	GCC	CCT	CCC
124.	FG5124	CGCGCGC	ACATG	ici	CIG	GGA	ATG	GCC	CCT	GCC .	CTG
		CGCGCGC	ACATG	TCT	GGA	ATG	ecc x10	CCT	GCC	CTG	CAG
125.	FGS125										
126.	FGS126	coccccc	ACATG	TCT	ATG	ecc	CCT	GCC	CIG	CAG	ccc
127.	FGS127	cececec	ACATG	TCT	GCC	CCT	GCC .	CIG	CAG	ccc	ACC
128.	FGS128	CGCCCCC	ACATG	TCT	CCI	GCC	CIG	CAG	CCC	ACC	CAG
129.	PGS129	CECECEC	<b>ACATG</b>	TCT	GCC	CIC	CAG	ccc	ACC	CAG	CCT
130.	PGS130	cececec	ACATG	TCT	CIG	CAG	CCC	ACC	CAG	CCT	GCC
131.	PGS131	CCCCCCC	ACATG	TCT	CAG	ccc	ACC	CAG	CGT	CCC	ATG
132.	FGS132	CGCGCGC	ACATG	TCT	CCC	ACC	CAG	CO.	GCC	ATG	· ccc
133.	FGS133	CGCGCGC	ACATG	TCT	ACC	CAG	GGT	GCC	ATG	ccs	GCC
134.	PG5134	CCCCCCC	ACATG	· ICI	CAG	GGT	GCC	ATG	-CCB	·ecc	TTC
135.	FGS135	CCCCCCC	ACATG	TCT	GGT	GCC	ATG	556	GCC	220	GCC
136.	FGS136	cececec	ACATG	TCT	GCC	ATG	ccc	GCC	TTC	GCC	TCT
137.		ececec	ACATG	TCT	ATG	CCC.	GCC		GCC	TCT	GCT
138.	FGS138	CGCGCGC	ACATG	TCT	ccc	GCC	TIC	GCC	TCT	GCT	TTC
139.	PG5139	CGCGCGC	ACATG	TCT	GCC	TTC	GCC	127	GCT	TTC	CAG
140.	PGS140	CCCCCCC	ACATG	TCT	TTC	GCC	TCT	622	TTC	CAG	CCC
141.	PGS141	cececec	ACATG	TCT	GCC.	TCT	GCT	TTT	CAG	CCC	CGG
	FGS142	CGCGCGC	ACATG	TCT	TCT	CCT	TTC	CAG	CGC	ccc	GCA
142. 143.		COCCCCC		101	GCT	770	CAG	252	-cec	CCA	GGA
	FGS143		ACATG								
244.	P35144	cacacac	ACATG	TCT	TTC	CAG	CCC	cea	GCA	GGA	GGG
145.	FGS145	coccccc	ACATG	TCT	CAG	CCC	CGG	SCY.	GGX	GGG	CIC
146.	FG5146	cececec	ACATG	TCT	CCC	CCC	GCX	GSX	GGG	GTC	CTG
147.	FGS147	cececes	ACATG	TCT	CCC	GCA	GGA	GCG	<b>GLC</b>	CTG	GTT
148.	FGS148	CCCCCCC	ACATG	TCT	GCA	GGX	GCG	C	CLC	GIT	GCT
149.	FG5149	CGCCCCC	ACATG	TCT	GGA	GGG	C1C	CTG	GII	GCT	AGC
150.	FGS150	CGCGCGC	ACATG	TCT	GGG	CTC	CTG	-Cii	-GCT	AGC	CAT
151.	FGS151	CCCCCCC	ACATO	TCT	GTC	CTG	CIT	GCT	AGC	CAT	CTG
252.	FG5152	CCCCCCC	ACATG	TCT	CTG	. GTT	GCT	ASC	CAT	CTG	CAG
151.	FGS153	cccccc	ACATG	TCT	CTT	CCT	AGC	CAT	CTG	CAG	AGC
254	PGS154	CCCCCCC	ACATG	TCT	GCT	AGC	CAT	623	CAG	AGC	TTC
255.	PGS155	COCCCCC	ACATG	TCT	AGC	CAT	CTG	CAG	AGC	TTC	CTG
154	FUS:56	2300000	ACATG	TCT	CAT	CIG.	CAG	ASS	110	CIG.	GAG
197.	FGS156	COCCCCC	ACATG	TCT	CTG	CAG	AGC	772	CTG	GAG	GTG
				•							
158.	PGS158	cccccc	ACATG	707	CAG	AGC	TTC	,077	GAG	GTG	TCG
159.		cccccc	ACATG	101	AGC	TTC	CIG	SAG	CTC	ICC	TAC
160.	PGS16C	SECREC	ACATG	TCT	TTC	CIG	GAG	GTG	TCS	TAC	CCC
161.	FG5161	2022032	ACATG	202	CIG	GAG	GTG	TES	TAC	·cs:	GTT
162.	FGS1E1	2300002	LIATS	121	- 3AG	GTG	TCG	TAT	cac	GTT	. CTA
163.	F35143	:306000	ACHTG.	207	223	TCG	TAC	535	CTT	CTA	csc
154.	PGS164	ದಾರಕರಕರ	*ことごこ	TCT	TCG	TAC	CGC	3	CTA	555	CAC
165.	PGS165	TREBEST.	ACATG	TCT	TAC	CGC	GTT	CTA	cac	CAC	C:73

166.	PGS166	cececec	ACATG	TET	CCC	CTI	CTA	·ccc	CAC	CTT	
167.	FGS167	CGCGCGC	ACATG	TCT	CII	CTA	CCC	CAC	Ct.	CCC	
168.	PGS168	CCCCCCC	ACATG	TCT	CTA	CCC	CAC	CIT	608	CAG	
169.	FGS169A	CCCCCCC	ACATG	TCT	CCC	CAC	ece c11	CAG	CCC	GA · C	
170.	FGS170A	CCCCCCC	ACATG .	TCT	CAC	ece C11	CAG	CCC	ey.c	ATG	
171.	FGS171A	cececec	ACATG ACATG	TCT	ece	CAG	CCC	CY.C	ATG	CCT	•
172.	FGS172A FGS173A	CCCCCCC	ACATG	TCT	CAG	CCC	GA'C	ATG	CCT	ACA	
173. 174.	FG5174A	CGCGCGC	ACATG	TCT	ccc	GY .C	ATG	GCT	ACA	CCA	
175.	FGS169B	CCCCCCC	ACATG	TCT	·ccc	CAC	CII	GCG	CAG	ccc	
176.	FGS170B	CGCGCGC	ACATG	TCT	CAC	CII	GCG	CAG	CCC	y.CI	
177.	FGS171B	CCCCCCC	ACATG	TCT	CII	CCC	CAG	ccc	Y.CL	AGT	
178.	FGS172B	CGCGCGC .	ACATG	TCI	GCG	CAG	ccc	A CT	AGT -CAT	CCA	
179.	FGS173B	CCCCCCC	ACATG	TCT	CCC	V.CI.	A ·CT	CAT	CCA	CCT	
180.	FGS174B	cececec	ACATG ACATG	TCT	CCC	CAC	CTT	GCG	CAG	CCC	
181.	FGS169C	CCCCCCC	ACATG	TCT	CAC	CTT	GCG	CAG	ccc	GGC	
182. 183.	FGS170C FGS171C	CECECEC	ACATG	TCT	CTT	GCG	CAG	CCC	GGC	CCC	
184.	FGS172C	CCCCCC	ACATG	TCT	GCG	CAG	CCC	GGC	GGC	GGC	
185.	FG5173C	CGCGCGC	ACATG	TCT	CAG	CCC	GGC	GGC	GGC	TCT	
186.	FGS174C	CCCCCC	ACATG	ICI	ccc	GGC	GGC	GGC	161	GA :C	
187.	RG50A	TATATAT	CCCCCCC	AGC	CAT	GTC	, GGG ACG	CTG .	CGC ACG	AAG ATT	
188.	RGSOB	TATATAT	ecceccec .	AGC AGC	CAT	GTC	AGA	GCC	GCC	GCC	
189.	RGSOC	TATATAT	ecceccec	TGT	AGC	CAT	GTC	GGG	CIG	CGC	
190.	RGS1A RGS1B	TATATAT	CCCCCCCCC	TOT	AGC	CAT	GTC	ACG	CCT	ACG	
191. 192.	RGS1C	TATATAT	CCCCCCC	TGT	AGC	CAT	CTC	AGA	GCC	GCC	
193.	RGS2A	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	CIC	GGG ,	CIG	
194.	RG52B	TATATAT	GCGGCCGC	TGG	TGT	AGC	CAT	esc.	ACG	CCT	
195.	RGS2C	TATATAT	CCCCCCCC	TGG	TGT	AGC	CAT	GTC	GTC	GCC	
196.	RGS3A	TATATAT	GCGGCCGC	CAA	TGG	TGT	AGC AGC	CAT	GTC	ACG	
197.	RGS3B	TATATAT	CCCCCCCCC	CAA	TGG	TGT	AGC	CAT	GTC	AGA	
198.	RGS3C RGS4	TATATAT TATATAT	GCGGCCGC	GCC	CYY	TGG	TGT	AGC	CAT	CTC .	
199. 200.	RGS5	TATATAT	CCCCCCC	AGG	GCC	CAA	TGG	TGT	AGC	CAT	
201.	RGS6	TATATAT	GCGGCCGC	GGC	AGG	GCC	CAA	TOG	TGT	AGC	
202.	RGS7	TATATAT	GCGGCCGC	GCT.	GCC	AGG	GCC	CAA	TGG	TCT	
203.	RGS8	TATATAT	ececcec	GGA	CCT	GGC	AGG GGC	GCC AGG	CYY	TGG CAA	
204.	RGS9	TATATAT	GCGGCCGC	CAG	GGA CAG	GCT GGA	CCT	GGC	AGG	GCC	
205.	RGS10	TATATAT	ececcec ececcec	CTG	GGG	CAG	GGY	GCT	GGC	AGG	
206. 207.	RGS11 RGS12	TATATAT TATATAT	GCGGCCGC	GCT	CTG	GGG	CAG	GGA	GCT	GGC	
207.	RGS13	TATATAT	CCCCCCC	GAA	GCT	CTG	GCC	CAG	GGA	GCT	
209.	RGS14	TATATAT	GCGGCCGC	CAG	GAA	GCT	CTG	GGG	CAG	GGA	
210.	RGS15	TATATAT	CCGCCCCC	GAG	CAG	GAA	CCT	CIG	GGG	CAG	
211.	RGS16	TATATAT	CCCCCCCC	CII	GAG	CAG	CYC	GCT GAA	GCT	CTG GGG	
212.	RGS17	TATATAT	GCGGCCGC	AGA TAA	CTT AGA	CTT	GAG	CAG	GAA	GCT.	
213.	RGS18	TATATAT TATATAT	CCCCCCCC	CIC	TAA	AGA	CII	GAG	CAG	GAA	
214.	RGS19 RGS20	TATATAT	GCGGCCGC	TTG	CTC	TAA	AGA	CII	GAG	CAG	
216.	RG521	TATATAT	GCGGCCGC	CAC	TIG	CIC	TAA	AGA	CII	GAG	
217.	RGS22	TATATAT	GCGGCCGC	CCT	CYC	TIG	CIC	TAA	AGA	CII	
218.	RGS23	TATATAT	CCCCCCCC	CII	CCT	CAC	TTG CAC	CTC	TAA CTC	AGA TAA	
219.		TATATAT	GCGGCCGC	GAT CTG	CTT	CLI	CCT	CAC	TIG	CTC	
220.	RGS25	TATATAT	GCGGCCGC	ecc erc	CTG	GAT	CII	CCI	CAC	TTG	
221.	RGS26 RGS27	TATATAT TATATAT	GCGGCCGC	ATC	GCC	CIG	GAT	CIT	CCT	CAC	
223.		TATATAT	GCGGCCGC	GCC	ATC	GCC	CIG	GAT	CTT	CCT	
224.		TATATAT	<b>GCGGCCGC</b>	TGC	GCC	ATC	GCC	CTG	GAT	CII	
225.		TATATAT	GCGGCCGC	CCC	TGC	GCC	ATC	GCC	CIG	GAT CTG	
226.		TATATAT	CCCCCCC	GAG	CGC	CCC .	GCC TGC	ATC GCC	GCC ATC	CIC	
227.		TATATAT	GCGGCCGC	CTC	CTG	GAG	CCC	TGC	GCC	ATC	
228. 229.		TATATAT TATATAT	GCGGCCGC	CIT	CIC	CTG	GAG	CCC	TGC	acc	
230.		TATATAT	GCGGCCGC	CAG	CTT	CTC.	CIG	GAG	CGC	TGC	
231		TATATAT	CCGCCCCC	ACA	CAG	CTT	CIC	CIG	GAG .	CCC	
232		TATATAT	SCGGCCGC	GGC	ACA	CAG	CIL	CIC	CZC	GAG	
233	. RGS38	TATATAT	CCCCCCC	CGT	GGC	ACA	CAG	CAG	CLC	CTC	
234		TATATAT	GCGGCCGC	GTA	GGT GTA	GGC GGT	. ACA	ACA	CAG	CIT	
235		TATATAT	60660060	CTT	CII	GTA	GGT	GGC	ACA	CAG	
.236		TATATAT TATATAT	GCGGCCGC	GCA	CAG	CII	GTA	GGT	GGC	ACA	
237. 239		TATATAT	GCGGCCGC	GTG	GCA	CAG	CTT	GTA	GGT	GCC	
239		TATATAT	eceeccec	GGG	GTG	GCA	CAG	CTT	GTA	CCT	
240		TATATAT	GCGGCCGC	CTC	GGG	GTG	GCA	CAG	CIT	GTA	
241	. RG\$46	TATATAT	eceeccec	CTC	CIC	GGG	GTG	GCA	CY3	CTI	
242		TATATAT	CCGCCCCC	CAG	CTC	CTC	GGG	GTG	GTG	GCA	
243		TATATAT	GCGGCCGC	CYC	CAG	CAG	CLC	CTC	GCS	GTG	
244		TATATAT	GCGGCCGC	CAG CAG	CAC	CAC	CAG	CTC	CTC	GGG	
245 246		TATATAT TATATAT	SCGGCCGC	. 100	GAG	CAG	CAC	CAG	CAG	CTC.	
247		TATATAT	SCGGCCGC	CIC	TCC	GAG	CAG	CAC	CAS	CTC	
248		TATATAT	GCGGCCGC "	AGA	CTG	TCC	GAG	CAG	CAC	CAG	
249		TATATAT	CCCCCCC	CAG	AGA	GTG	TCC	GAG	EAG	CAC	
250		TATATAT		ecc	CAG	AGA	C1C	355	GAS	CAS	
251	. RGES6	TATATAT	<b>CCGCCCCC</b>	-CAT	GCC	CAG	AGA	CTG	755	GAG -	

252.	RGS57	TATATAT	CCCCCCCC	GGG	GAT	GCC	CAG	ASA	GEG	TCC
252. 253.	RGS58	TATATAT	GCGGCCGC	<b>√CC</b> λ	CCC	SAT	GCC	CAG	AGA	GTG
254.	RGS59	TATATAT	CCCCCCCC	AGC	CCX	CCC	GAT	GCC	CAG	AGA
255.	RGS60	TATATAT	CCCCCCCC	GGG	AGC	CCA	GGG	-GAT	GCC	CXC
256.	RGS61	TATATAT	CCGCCCCC	CAG	COG	AGC	CCA	CCC	GAT	ccc
257.	RGS62	TATATAT	GCGGCCGC	CCT	CAG	GGC	AGC	·CCA	GGG	GAT
258.	RGS63	TATATAT	CCCCCCC	GGA	CCT	CAG	CCC	AGC	CCA	GCG
259.	RGS64	TATATAT	CCCCCCCC	GCA	GGA	CCT	CAG	.ccc	AGC	CCY
260.	RGS65	TATATAT	CCGCCCCC	GGG	GCA	CCA	GCT	CAG	GGG	AGC
261.	RGS66	TATATAT	CCCCCCC	GCT	GGG	GCA	GGA	CCT	CAG	GGG
262.	RGS67	TATATAT	CCCCCCCC	CTG	GCT	GGG	GCA	GCX	GCT	CAG
263.	RGS68	TATATAT	GCGGCCGC	GCC	CTG	GCT	GGG	-GCA	GGA	GCT
264.	RGS69	TATATAT	CCCCCCCC	CAG	CCC	CTG	GCT	CCC	GCA	GGA
265.	RG570	TATATAT	CCCCCCCC	CTG	CAG	GGC	CIG	CCT	GGG	GCA
266.	RG571	TATATAT	<b>CCCCCCC</b>	CAG	CTG	CAG	GGC	CTG	GCT	GGG
267.	RGS72	TATATAT	CCCCCCCC	TGC	CAG	CTG	CAG	GGC	CIG .	GCT
268.	RGS73	TATATAT	CCCCCCC	CCC	TGC	CAG	CIG	CAG	GGC .	C1C
269.	RGS74	TATATAT	GCGGCCGC	GCA	GCC	TGC	CAG	CIG	CAG	GCC
270.	RGS75	TATATAT	<b>GCGGCCGC</b>	CAA '	GCA	GCC	TGC	CAG	CIG	CAG
271.	RGS76	TATATAT	<b>CCGGCCGC</b>	GCT	CAA	GCX	GCC	TGC	CXG	.crg
272.	RGS77	TATATAT	ccccccc	GII	CCT	CYY	GCA	GCC	TGC	TGC
273.	RGS78	TATATAT	CCCCCCCC	GAG	CIT	GCT	CAA	GCX	-GCC	GCC 1GC
274.	. RGS79	TATATAT	CCCCCCC	ATG	GAG	CTT	GCT	CXX	GCX	GCA
275.	RGSBO	TATATAT	CCCCCCCC	GCT	YIC	GAG	GTT	GCT	CYY	CAA
276.	RGS81	TATATAT	CCCCCCCC	CCC	CCT	ATG	GAG	GTT	GCT	GCT
277.	RGS82	TATATAT	CCGCCCCC	AAG	GCC	CCT	ATG	GAG	GTT	GTT
278.	RGS83	TATATAT	GCGGCCGC	GAA	YYC	GCC	CCT	ATG	GAG ATG	GAG
279:	RG584	TATATAT	ececcec	GAG	GYY	N.G	GCC	GCC	GCT	ATG
280.	RG585	TATATAT	CCCCCCCC	GTA	GAG	GAA	AAG	AAG	GCC GC1	GCT
281.	RGS86	TATATAT	CCCCCCCC	CTG	GTA .	GAG	GAX	GAA	AAG	GCC
282.	RG587	TATATAT	CCCCCCCC	CCC	CIG	GTA CTG	GAG GTA	GAG	GAA	AAG
283.	RGS88	TATATAT	GCGGCCGC	GAG	ecc	CCC	CIG	GTA	GAG	GAA
284.	RGS89	TATATAT	CCCCCCC	CAG	GAG CAG	GAG	CCC	CTG	GTA	GAG
285.	RGS90	TATATAT	GCGGCCGC	GGC CTG	CAC	CAG	GAG	CCC	CTG	GTA
286.	RGS91	TATATAT	GCGGCCGC	CAG	GGC	CIG	CAG	GAG	ccc	CTG
287.	RGS92	TATATAT	GCGGCCGC.	TIC	CAG	CCC	CIG	CAG	GAG	·ccc
288.	RGS93	TATATAT	GCGGCCGC	CCC	TTC	CAG	GGC	CIG	CAG	GAG
289.	RGS94	TATATAT	GCGGCCGC	TAT	ccc	TIC	CAG	GGC	CTG	CAG
290.		TATATAT	GCGGCCGC GCGGCCGC	CCY	TAT	CCC	TTC	CAG	GGC	-CTG
291.	RGS96	TATATAT	GCGGCCGC	GGG	GGA	TAT	CCC	TTC	CAG	GGC
292.	RG597	TATATAT	GCGGCCGC	CIC	GCC	GGA	TAT	-000	TTC	CAG
293.	RGS98	TATATAT TATATAT	GCGGCCGC	CAA	CTC	GGG	GGA	TAT	CCC	TIC
294.	RGS99	TATATAT	GCGGCCGC	ACC	CAA	CTC	GGG	GGA	TAT	ccc
295.	RGS100	TATATAT	CCCCCCC	GGG	ACC	CAA	CTC	GGG	GGA	TAT
296.	RGS101 RGS102	TATATAT	CCGCCCCC	GGT	GGG	ACC	CAA	CTC	GGG	CCY
297.	RGS103	TATATAT	CCCCCCC	CAA	GGT	GGG	ACC	CXX	CIC	CCC
298.		TATATAT	GCGGCCGC	GTC	CAA	GGT	GGG	ACC	CAA	CTC
299. 300.		TATATAT	GCGGCCGC	TGT	GTC	CAA	GGT	GGG	ACC	CAA
300.		TATATAT	GCGGCCGC	CAG	TCT	CIC	CAA	GUT	GGG	ACC
302.		TATATAT	CCGCCCCC	CTG	CAG	TGT	GTC	CAA	GGT	GGG
303.		TATATAT	GCGGCCGC	CAG	CTG	CAG	TCT	<b>GTC</b>	CAA	GCT
304.		TATATAT	GCGGCCGC	GTC '	CAG	CIG	CAG	TOT	*C10	CYY
305.		TATATAT	GCGGCCGC	GAC	GTC	CAG	CIG	CXC	TGT	GTC
306	The second secon	TATATAT	GCGGCCGC	GGC	GAC	GTC	CAG	CIG	CAG	TGT
307		TATATAT	GCGGCCGC	GTC	GGC	GAC	GTC -	CAG	CTG	CAG
308		TATATAT	CCCCCCC	AAA	GIC	GGC	GAC	CIC	·CAG	CTG
309		TATATAT	GCGGCCGC	GGC	AAA	GTC	GGC	GAC	GTC	CAG
310		TATATAT	GCGGCCGC	CCT	CCC	XXX	CIC	GGC	GAC	CTC
311		TATATAT	CCGGCCCC	GGT	CCT	GGC	YYY	CLC.	GGC	GAC
312		TATATAT	CCCCCCCC	GAT	CCT	CCT	CCC	***	GTC	GGC
313	. RGS118	TATATAT	CCCCCCCC	CCX	GAT	GGT	CCT	GGC	GGC	AAA
314	. RGS119	TATATAT	ecaecaec	CTG	CCX	GAT	GGT	GGT	GGT	GGC
315	. RGS120	TATATAT	GCGCCCCC	cte	CTG	CCY	GAT	GGT	GCT	GGT
316	. RG5121	TATATAT	eceeceec	CAT	CIG	CTG	CCY	-GAT CCA	GAT	GGT
317	. RGS122	TATATAT	CCCCCCC	110	CAT	CIG	CIG	CIG	CCY	GAT
318	. RGS123	TATATAT	ececcec	TIC	170	CAT	CTG	·C1G	CTG	CCY
319	. RGS124	TATATAT	GCGGCCGC	CAG	11C	TTC		CAT	CTG	CTG
320	. RGS125	TATATAT	GCGCCCCC	TCC	CXG	TTC	TTC	11C	CAT	CTG
321	RGS126	TATATAT	GCGGCCGC	CAT	TCC	TCC	CAG	TTC	TTC	CAT
322	. RG5127	TATATAT	GCGGCCGC	GGC	CAT		TCC	-CAG	110	TTC
323	. RGS128	TATATAT	CCGCCCCC	AGG	GGC	CAT	CAT	TCC	CAG	TTC
324		TATATAT	GCGGCCGC	CCC		GGC AGG	· eec	CAT	100	CAG
325		TATATAT	GCGGCCGC	CYC	GGC	GGC	AGG	CVI	CAT	202
326		TATATAT	ececcec	CLO	CAG		GGC	AGG	GGC	CAT
323		TATATAT	GCGGCCGC	GGG	CLC	CAG	CAG	GGC	- AGG	GGC
125		TATATAT	GCGGCCGC	GGT	GGG	CLC	CIG	CAG	GCC	AGG
329		TATATAT	GCGGCCGC	. CTG	COT	GGG	CIG	CTG	CAG	GSC
330		TATATAT	coccccc	YCC	CTG	CCI	GCT	CCC	C.c.c	CAS
33:		TATATAT	GOGGCCGC	GGC	YCC	CTG ACC	CTG	CCT	653	C12
33		TATATAT	GCGGCCGC	CAT	GGC	GGC	ACC	CTG.	GUT	653
33		TATATAT	GCGGCCGC	CGC	CAT	CAT	GGC	YEC	cre	GGT
. 334		TATATAT	GCGGCCGC	GGC	ccc	CCC	CAT	GGC	ATT	cre
33		TATATAT	GCGGCCGC			GGC	CCC	CAT	900	YEE
2 2 1		TATATAT	GCGGCCGC	GGC	GAA		. eec	-CGG	CAT	360
33	7. RGS142	TATATAT	-CCGCCCCC	AGA	GGC	SAA	666	.00	٠	200

338.	RGS143	TATATAT	GCGGCCGC	AGC	AGA	GGC	GAA	GGC	CGG	CAT	
339.	RG5144	TATATAT	GCGGCCGC	GAA	AGC	AGA	GGC	GAA	ccc	CCC	
340.	RGS145	TATATAT	GCGGCCGC	CTG	GAA	AGC	AGA	GGC	GAA	CGC	
341.	RGS145	TATATAT	CCCCCCCC	GCG	CTG	GAA	AGC	AGA	GCC	GAA	
342.	RGS147	TATATAT	GCGGCCGC	CCC	GCG	CTG	GAA	AGC	AGA	CCC	
	RGS148	TATATAT	GCGGCCGC	TGC	CCG	·GCG	CTG	GAA	AGC	AGA	-
343. 344.	RGS149	TATATAT	GCGGCCGC	TCC	TGC	CCG	GCG	CTG	GAA	AGC	•
345.	RGS150	TATATAT	CCGCCCCC	ccc	TCC	TGC	ccc	CCC	CTG	GAA	
345.	RGS150	TATATAT	GCGGCCGC	GAC	ccc	TCC	<b>7GC</b>	CCC	ccc	CTG	
345.	RGS152	TATATAT	GCGGCCGC	CAG	GAC	ccc	TCC	TGC	CCG	GCG	
-	RG5153	TATATAT	GCGGCCGC	AAC	CAG	GAC	-000	TCC	TGC	ccc	
348. 349.	RG5154	TATATAT	GCGGCCGC	AGC	AAC	CAG	GAC	ccc	TCC	TGC	
	RGS155	TATATAT	CCCCCCCC	GCT	AGC	AAC	CAG	GAC	CCC	TCC	
350.	RGS156	TATATAT	CCCCCCC	ATG	CCT	AGC	AAC	CAG	GAC	ccc	
351.	RGS150	TATATAT	GCGGCCGC	CAG	ATG	GCT	AGC	AAC	CAG	GAC	
352.	RGS158	TATATAT	CCCCCCCC	CTG	CAG	ATG	GCT	AGC	AAC	CAG	
353.	RG5159	TATATAT	CCCCCCC	CCT	CTG	CAG	ATG	GCT	AGC	AAC	-
354.	RG5159	TATATAT	eceecee	GAA	CCT	CTG	CAG	ATG	GCT	AGC	
355. 356.	RGS161	TATATAT	GCGGCCGC	CAG	GAA	GCT	CIG	CAG.	ATG	CCT	
	RGS162	TATATAT	GCGGCCGC	CTC	CAG	GAA	GCT	CIG	CXG	ATG	
357.	RGS163	TATATAT	GCGGCCGC	CAC	CTC	CAG	GAA	GCT	CIG	CAG	
358. 359.	RGS164	TATATAT	CCCCCCCC	CGA	CAC	CIC	CAG	GAA	GCT	CIG	
	RGS165	TATATAT	GCGGCCGC	GTA	CGA	CAC	CTC	CAG	GAA	CCT	
360. 361.	RGS166	TATATAT	GCGGCCGC	GCG	GTA	CGA	CAC	CTC	CAG	GAA	
362.	RGS167	TATATAT	- GCGGCCGC	AAC	GCG	GTA	CGA	CAC	CTC	CAG	
363.	RGS168	TATATAT	GCGGCCGC	TAG	AAC	CCC	GTA	CGA	CAC	CTC	•
364.	RGS169	TATATAT	CCGCCCCC	GCG	TAG	AAC	CCC	GTA	CGA	CAC	
365.	RGS170	TATATAT	GCGGCCGC	GTG	GCG	TAG	AAC	GCG	GTA	CGA	
366.	RGS171	TATATAT	CCCCCCC	AAG	GTG	GCG	TAG	AAC	GCG	GTA	
367.	RGS172	TATATAT	GCGGCCGC	CGC	AAC	GTG	GCG	TAG	AAC-	·ece	
367.	PGS173	TATATAT	ccccccc	CTG	cac	AAG	GTG	GCG	TAG	AAC	GCG

10

15

20

#### **CLAIMS**

#### What is claimed is:

1. A method for making a biologically-active circularly-permuted protein of the formula C¹-L¹-N¹, derived from a parent protein of the formula N¹-C¹, wherein

C¹ is comprised of a segment derived from the carboxy portion of said parent protein;

N¹ is comprised of a segment derived from the amino terminal portion of said parent protein; and

 $L^1$  is a chemical bond or a linker, linking  $C^1$  to the amino terminus of  $L^1$  and carboxy terminus of  $L^1$  to the amino terminus of  $N^1$ ;

comprising the steps of:

- (a) making a series of circularly-permuted genes;
- (b) inserting said circularly-permuted genes into a display vector;
- expressing said circularly-permuted genes such that the proteins encoded by said genes are presented on the surface of the display vector;
- (d) generating a library of display vectors presenting the expressed circularly permuted protein;
- (e) affinity-selecting the presenting display vectors with a target protein that can bind a biologically-active circularly-permuted protein;
- (f) isolating and analyzing clones of selected display vectors to identify the presented circularly-permuted protein.
- 25 2. The method of claim 1 wherein the method of making a series of circularly-permuted genes is selected from the group consisting of making a tandemly-repeated intermediate, total synthesis of a synthetic gene, assembly of a gene from synthetic oligonucleotides, DNA amplification, and limited digestion of a circular

intermediate.

- 3. The method of claim 1 wherein said display vector is selected from the group consisting of bacteriophage display vectors, bacteria, and baculovirus vectors.
  - 4. The method of claim 3, wherein said presentation vector is a bacteriophage.
    - The method of claim 4, wherein said presentation vector is bacteriophage M13.
      - 6. The method of claim 5, wherein said presentation vector is a bacteriophage M13 gene III vector.
- 7. The method of claim 1 wherein said method of making a series of circularly permuted genes is a method of making a tandem repeat intermediate.
  - The method of claim 7 wherein said circularly-permuted genes are amplified from the repeat by gene amplification.
- 9. The method of claim 1 wherein said method of affinity selection comprises the steps consisting of:
  - (a) binding said presentation display vectors to a target protein;
  - (b) eluting said display vectors;
  - (c) amplifying said display vectors; and
  - (d) biopanning a pool of said amplified display vectors.
- 10. The method of claim 1 wherein L<sup>1</sup> is a linear peptide linker.
- 11. The method of claim 1 wherein said the DNA sequence encoding said linker L<sup>1</sup> is selected from the group consisting of SEQ ID NO: 1 through SEQ ID NO: 368.
- 12. The method of claim 1 wherein the length of the  $C^1$  in said permutein is shorter than the length of  $C^1$  in said parent protein.

10

5

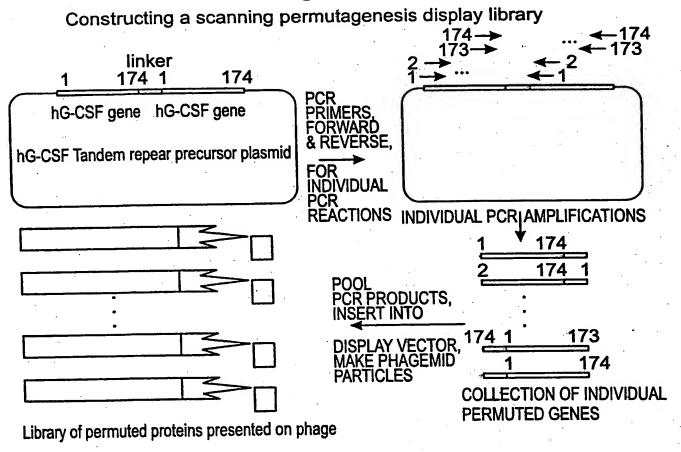
15

20

- 13. The method of claim 1 wherein the length of the  $N^1$  in said permutein is shorter than the length of  $N^1$  in said parent protein.
- 14. A circularly-permuted protein prepared by the method of claim 1.
  - 15. A circularly-permuted protein of claim 14 comprising the G-CSF receptor agonist domain of a species of mylepoietin (MPO).

1/4

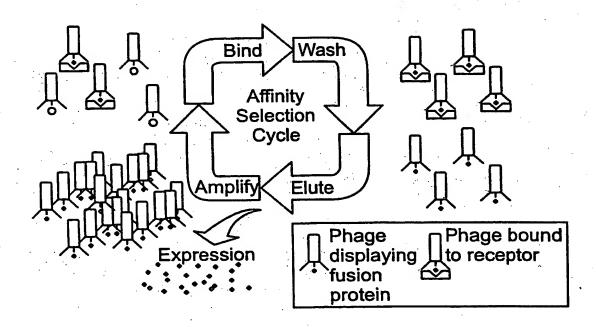
Figure 1A



2/4

# Figure 1B

Screening a display library



3/4

Figure 2

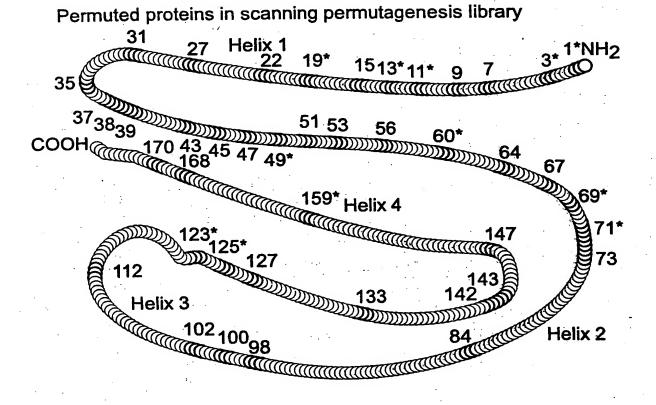
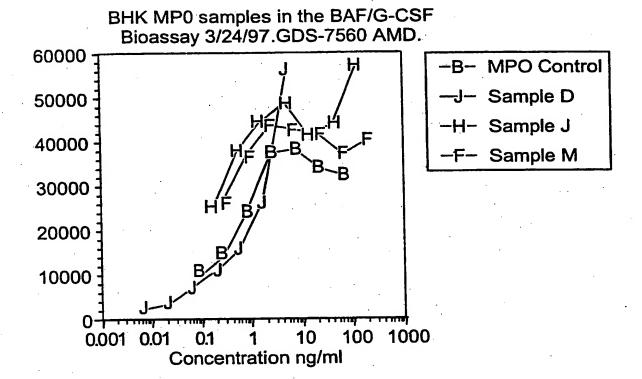


Figure 3



## SEQLIST.TXT

### SEQUENCE LISTING

<110> Lee, S	tephen C					
	• •					
	of Producing Per ermutagenesis	rmuteins by				
<130> C_3041						
<140> 60/	,			•		
<141> 1998-0	9-25			٠		
<160> 368						
<170> FastSE	Q for Windows Ve	rsion 3.0				
<210> 1	•				• -	
<211> 36			•	:		
<212> DNA		•	•			
<213> Synthe	etic					
<400> 1						
	acc attgggccct g	ccagc				36
<210> 2		-				
<211> 36						
<212> DNA						
<213> Synthe	etic					
<400> 2	$\cdot$	٠.	*			
	catt gggccctgcc a	gctcc				3 (
-2105 3						
<210> 3 <211> 36		•				. 0
<211> 30 <212> DNA				-		
<213> Synthe	etic	,	. :			
	•					
<400> 3	tggg ccctgccagc t	cccra	•			3 (
gcgcgcaca cgcccc			•			
<210> 4					•	
<211> 36						
<212> DNA						
<213> Synthe	etic			•		
<400> 4						_
egegegeaca tgtetg	gccc tgccagctcc (	etgeee				31
<210> 5						
<211> 3£	•		•			
<212> DNA						
<213> Synth	etic					
<400> 5						
cgcgcgcaca tgtctc	ctgc cagctccctg o	cccag				3
<210> 6						
<211> 36	•				•	
<212> DNA						•
<213> Synth	etic	*				
<40C> 6						
	ecaa etecetacee «	cagage			•	3

<210> 7		
<211> 36		
<212> DNA		
<213> Synthetic		
·		
<400> 7		
cgcgcgcaca tgtctagctc cctgccccag agcttc		36
<210> 8		
<211> 36	•	
<212> DNA		
<213> Synthetic	· ·	
<400> 8		
cgcgcgcaca tgtcttccct gccccagagc ttcctg		36
<210> 9	•	
<211> 36		
<212> DNA	•	
<213> Synthetic		
	•	
<400> 9		36
cgcgcgcaca tgtctctgcc ccagagette ctgctc	•	36
<210> 10		
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 10		36
cgcgcgcaca tgtctcccca gagcttcctg ctcaag	·	30
<210> 11		
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 11		36
cgcgcgcaca tgtctcagag cttcctgctc aagtct		50
<210> 12		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 12		36
cgcgcgcaca tgtctagctt cctgctcaag tcttta	•	
<210> 13		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 13		. 36
egegegeaca tgtettteet geteaagtet ttagag	·	
<210> 14		
•		
<211> 36		
<212> DNA		
<213> Synthetic		
-400> 14		
<400> 14		36
cgcgcgcaca tgtctctgct caagtcttta gagcaa	•	
-010x 35		
<210> 15		
<211> 36		
<211> 3c <212> DNA <213> Synthetic		

<210> 24

<400> 15 cgcgcgcaca tgtctctcaa	gretttagag	caagtg		36
<210> 16				
<211> 36				
<212> DNA			·	
<213> Synthetic		•		
<400> 16				3.6
cgcgcgcaca tgtctaagtc	tttagagcaa	gtgagg		36
<210> 17				
<211> 36			•	
<212> DNA	•			
<213> Synthetic			•	
<400> 17			•	36
cgcgcgcaca tgtcttcttt	agagcaagtg	aggaag		36
<210> 18	·	.*		
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 18				
cgcgcgcaca tgtctttaga	gcaagtgagg	aagatc		3.6
<210> 19				
<211> 36	•	•	•	•
<212> DNA	•		•	•
<213> Synthetic			*	
<400> 19				*
cgcgcgcaca tgtctgagca	agtgaggaag	atccag	•	36
				•
<210> 20				
<211> 36			•	•
<212> DNA				
<213> Synthetic	i		. •	
<400> 20			•	36
cgcgcgcaca tgtctcaagt	gaggaagatc	cagggc		30
<210> 21				
<211> 36	•		*	•
<212> DNA			. •	
<213> Synthetic	: ,		. *	
<400> 21				
cgcgcgcaca tgtctgtgag	gaagatccag	ggcgat	•	36
<210> 22				
<211> 36				
<212> DNA				
<213> Synthetic	•			
<400> 22			·	
cgcgcgcaca tgtctaggas	gatccagggc	gatggc		36
<210> 23	•		· *	
<211> 36		•		
<212> DNA				
<213> Synthetic	:			
-100× 22		•		
<400> 23 cgcgcgcaca tgtctaagat	ccagggcgat	ggcgca	·	36
			•	

<211> 36 <212> DNA					
<213> Synthetic				•	•
<400> 24					
cgcgcgcaca tgtctatcca	gggcgatggc	gcagcg			36
<210> 25		٠			
<211> 36 <212> DNA	·				
<213> Synthetic					
<400> 25					
cgcgcgcaca tgtctcaggg	cgatggcgca	gcgctc		3	3 (
<210> 26			•		
<211> 36				•	
<212> DNA <213> Synthetic	•				
<400> 26					
cgcgcgcaca tgtctggcga	tggcgcagcg	ctccag			36
<210> 27					
<211> 36	•	-	•	•	
<212> DNA <213> Synthetic	:			•	
<400> 27	. •		•		
cgcgcgcaca tgtctgatgg	cgcagcgctc	caggag			36
<210> 28		•			
<211> 36					
<212> DNA <213> Synthetic					
vals, symmetric		•			•
<400> 28 cgcgcgcaca tgtctggcgc	agegetecag	gagaag			36
***					
<210> 29 <211> 36			•	•	
<211> 30 <212> DNA					
<213> Synthetic	:				
<400> 29					
cgcgcgcaca tgtctgcago	gctccaggag	aagctg			3 (
<210> 30			•		
<211> 36			•		
<212> DNA			•		
<213> Synthetic	•				
<400> 30		. statat			. 3 (
cgcgcgcaca tgtctgcgct	. ccayyayaay	ctgtgt	•		
<210> 31		•			٠.
<211> 36					
<212> DNA <213> Synthetic	<b>-</b> .				
	-				
<400> 31					
egegegeaca tgtetetee	a ggagaagctg	tgtgcc			36
<210> 32			•	•	
<211> 36 <212> DNA			• •		
<213> Synthetic	c		. •		

egegegeaca tgteteagga gaagetgtgt gecaee	30
<210> 33	
<211> 36	
<212> DNA	-
<213> Synthetic	
<400> 33	
cgcgcgcaca tgtctgagaa gctgtgtgcc acctac	36
<210> 34	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 34	
cgcgcgcaca tgtctaagct gtgtgccacc tacaag	36
<210> 35	
<211> 36	
<212> DNA	
<213> Synthetic	
400. 20	
<400> 35 cgcgcgcaca tgtctctgtg tgccacctac aagctg	36
cgcgcgcaca tgtctctgtg tgctattat anguly	
<210> 36	٠.
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 36	36
egegegeaca tgtettgtge cacetacaag etgtge	
<210> 37	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 37	36
cgcgcgcaca tgtctgccac ctacaagctg tgccac	
<210> 38	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 38	36
cgcgcgcaca tgtctaccta caagctgtgc cacccc	
<210> 39	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 39	36
cgcgcgcaca tgtcttacaa gctgtgccac cccgag	•
<210> 40	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 40	
cgcgcgcaca tgtctaagct gtgccacccc gaggag	36
2010- 41	
<210> 41	
<211> 36	

ZIII Synchecte		
<400> 41		
cgcgcgcaca tgtctctgtg ccaccccgag	gagetg	36
<210> 42		
<211> 36		
<212> DNA		
<213> Synthetic		•
-400- 42		
<400> 42 cgcgcgcaca tgtcttgcca ccccgaggag	ctggtg	36
cdcdcdcaca cdcccdcca cccode32323		
<210> 43		
<211> 36		
<212> DNA		
<213> Synthetic		
<pre>&lt;400&gt; 43 cgcgcgcaca tgtctcaccc cgaggagctg</pre>	gracta	36
Egegegeaca egecteace egaggageeg	969629	
<210> 44	• •	
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 44	oract o	36
cgcgcgcaca tgtctcccga ggagctggtg	etgete	
<210> 45		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 45	et eggs	36
cgcgcgcaca tgtctgagga gctggtgctg	ctcgga	
<210> 46	·	
<211> 36		
<212> DNA		
<213> Synthetic		
*		
<400> 46	GG2C2C	36
cgcgcgcaca tgtctgagct ggtgctgctc	ggacac	-
<210> 47	•	
<211> 36		
<212> DNA	•	
<213> Synthetic		
-400- 47		
<400> 47 cgcgcgcaca tgtctctggt gctgctcgga	cactet	36
cycycycaca tytotocygo yorgettyya		
<210> 48	·	
<211> 36	·	
<212> DNA		٠.
<213> Synthetic	•	
.40040		
<400> 48	teteta	36
cgcgcgcaca tgtctgtgct gctcggacac	coccin	30
<210> 49		
<211> 36	•	
<212> DNA	*	
<213> Synthetic	•	
•		
<400> 49		
COCOCOCACA FORCECTOCT COGACACTCI	ctagac	36

<210> 50						
<211> 36						
<212> DNA						
<213> Synthetic						
C213> Synthetic						
<400> 50						36
cgcgcgcaca tgtctctcgg a	acactctctg g	gcatc				30
<210> 51			•			
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 51			•			
cgcgcgcaca tgtctggaca	ctctctgggc a	teccc	· .			36
<210> 52						
<211> 36						
<212> DNA						
<213> Synthetic	•					
<400> 52						
cgcgcgcaca tgtctcactc	tetgggcate c	cctgg				36
•	•	.:		•		
<210> 53						
<211> 36 <212> DNA						
<213> Synthetic						
<400> 53						36
cgcgcgcaca tgtcttctct	gggcatecee t	gggct				30
<210> 54					* •	
<211> 36						
<212> DNA	·		•*			
<213> Synthetic			÷			
<400> 54 cgcgcgcaca tgtctctggg	esteccetas c	retece				36
cacacaca raccerada	Carecees 5	,0000				
<210> 55	•					
<211> 36						
<212> DNA	• •					
<213> Synthetic				•		
<400> 55	•					
cgcgcgcaca tgtctggcat	cccctgggct	ecctg				36
<i>-</i>						
<210> 56				•		
<211> 36						
<212> DNA						
<213> Synthetic						
<400> 56						
cgcgcgcaca tgtctatccc	ctgggctccc	ctgagc .		•		36
<210> 57						
<211> 36						
<212> DNA		•				
<213> Synthetic	•					
<400> 57	•					
egegegeaca tgteteectg	ggctcccctg	agctcc				36
-010- 60			•	٠		
<210> 58			•		•	
<211> 36						
<212> DNA <213> Synthetic	•		-		-	••
<\$13> Synchecic	•					

<400> 58						
cgcgcgcaca tgtcttgggc	tcccctgagc	tectge				36
						•
<210> 59						
<211> 36						
<212> DNA						
<213> Synthetic				•		
,						
<400> 59			•			
cgcgcgcaca tgtctgctcc	cctgagctcc	tgcccc				36
<210> 60						
<211> 36						
<212> DNA						
<213> Synthetic	•					
<400> 60						
cgcgcgcaca tgtctcccct	anactertae	CCCBGC				3.6
egegegeaca egecteete	gageteetge	cccage				
<210> 61	-					
<211> 36						
<212> DNA						
<213> Synthetic						
				•		
<400> 61						
cgcgcgcaca tgtctctgag	ctcctgcccc	agccag		•	•	36
<210> 62						
<211> 36						
<212> DNA						
<213> Synthetic	:			•		
<400> 62		CREECO				36
egegegeaca tgtctagete	ccycccayc	caggee				30
<210> 63						
<211> 36						
<212> DNA	•					
<213> Synthetic	:					
<400> 63		•				
cgcgcgcaca tgtcttcctg	ccccagccag	gccctg				36
_						
<210> 64				•		
<211> 36						
<212> DNA	_		•	* .		
<213> Synthetic	i					
<400> 64						
cgcgcgcaca tgtcttgccc	CAGCCAGGCC	ctocao				36
-202-2		1 -23				
<210> 65	•					
<211> 36						
<212> DNA						
<213> Synthetic	2					
						•
<400> 65			•			
cgcgcgcaca tgtctccca	g ccaggccctg	cagetg				36
<210> 66						
<211> 36						
<212> DNA			•			
<213> Synthetic	-					
<400> 66						
cgcgcgcaca tgtctagcc	, a ggccctgcad	Ctaaca				36
-9-2-25	. 3333-1	3 3				
<210> 67		•				
<211> 36						
<211> 36						

•						
<212> DNA						
<213> Synthetic						
•						
<400> 67						
cgcgcgcaca tgtctcaggc	. cctacaacta	acadac				36
cycycycaca cyceccagge		9				
4710 60						
<210> 68						
<211> 36			•			
<212> DNA	•					
<213> Synthetic	•			•		
. <400> 68						2.0
cgcgcgcaca tgtctgccct	. gcagctggca	ggctgc	4			36
				•		
<210> 69			•			
<211> 36	1					
<212> DNA						
<213> Synthetic	:	-				
<400> 69	Δ	·				
cgcgcgcaca tgtctctgca	gctggcaggc	tgcttg				36
<210> 70					••	
<211> 36						
<212> DNA	:		•		+ .	
<213> Synthetic	<u>.</u>					
	-					
<400> 70						
cgcgcgcaca tgtctcagc	ggcaggctgc	ttgage				36
cgcgcgcaaa cgtccc	- 55555					
<210> 71						•
<211> 36				•		
<212> DNA						
<213> Synthetic	_					
<213> Synchectic				•	•	
<400> 71						
cgcgcgcaca tgtctctgg	c agactacttc	anccaa				36
cgcgcgcaca cgcccccgg	c aggergere	,		*		
<210> 72					•	
<211> 36						
<212> DNA	•					
	^					
<213> Synthetic						
<400> 72						
cgcgcgcaca tgtctgcag	a cracttaga	CRACTC				36
egegegeaca egeetgeag	y cryceryay.	Cuucu		•,		٠
<210> 73						
<211> 36 <212> DNA						
<213> Syntheti	_					
12132 Syntheti	_			•		
<400> 73						
cacacaca tatotaget	~ ~ C	rtccat				36
-Achedeaca threedaca	a ceranacen					
<210> 74					•	
<211> 36						
<212> DNA	_					
<213> Syntheti	C					
			•			
<400> 74						36
cgcgcgcaca tgtcttgct	t gagccaact	catage				30
		•				
<210> 75						
<211> 36	_		•			
<212> DNA	•			٠.		
<213> Syntheti	c					
<400> 75						
						7.4

<210× 70		
<211> 36		
<212> DNA		
<213> Synthetic		
•	•	
<400> 76		36
gegegeaca tgtetageca actecatage ggeett		20
<210> 77		
<211> 36		
<212> DNA		
<213> Synthetic		
•		
<400> 77		36
cgcgcgcaca tgtctcaact ccatagcggc cttttc		30
<210> 78		
<211> 36		
<212> DNA		
<213> Synthetic		
50		
<400> 78		36
egegegeaca tgteteteca tageggeett tteete		
-210> 70	, , , , , , , , , , , , , , , , , , ,	
<210> 79	•	
<211> 36 <212> DNA		
<212> DNA <213> Synthetic		
2213> Synthetic		
<400> 79		
cgcgcgcaca tgtctcatag cggccttttc ctctac		36
egegegezea egeeedadag eggeeeet		•
<210> 80		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 80	•	•
cgcgcgcaca tgtctagcgg ccttttcctc taccag		36
<210> 81		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 81		
cgcgcgcaca tgtctggcct tttcctctac cagggg		36
<210> 82		
<211> 36	•	
<212> DNA		
<213> Synthetic		
	•	
<400> 82		. 36
cgcgciceca tgtctctttt cctctaccag gggctc	• .	
<210> 83		
<211> 36		
<212> DNA		
<213> Synthetic	•	
	•	
<400> 83		3 (
cgcgcgcasa tgtcsttcct ctaccagggg ctcctg		
<210> 84		
<211> 36	•	
<212> DNA		
<2:3> Symthetic		

<400> 84		36
cgcgcgcaca tgtctctcta ccaggggctc ctgcag		
<210> 85		
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 85		36
cgcgcgcaca tgtcttacca ggggctcctg caggcc		30
<210> 86		
<211> 36	•	
<212> DNA		
<213> Synthetic		
<400> 86	•	36
cgcgcgcaca tgtctcaggg gctcctgcag gccctg		
<210> 87		
<211> 36	Ġ:	
<212> DNA	· :	
<213> Synthetic	•	
<400> 87		36
egegegeaca tgtetggget cetgeaggee etggaa	•	.30
<210> 88		
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 88		
cgcgcgcaca tgtctctcct gcaggccctg gaaggg		36
<210> 89		
<211> 36		•
<212> DNA		
<213> Synthetic		
<400> 89	·	36
cgcgcgcaca tgtctctgca ggccctggaa gggata		36
<210> 90	•	
<211> 36		
<212> DNA		
<213> Synthetic	•	
<400> 90		36
cgcgcgcaca tgtctcaggc cctggaaggg atatcc		50
<21%> 9?		
<211> 36	•	
<212> DNA	• •	
<213> Synthetic		
<400> 91		26
egegegeaca tgretgeeer ggaagggata teeece		36
<210> 92		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 92		
cgcgcgcaca tgtctctgga agggatatcc cccgag		36

<210> 93

<211> 36	
<212> DNA	
<213> Synthetic	
400. 02	
<400> 93	36
cgcgcgcaca tgtctgaagg gatatccccc gagttg	50
<210> 94	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 94	
cgcgcgcaca tgtctgggat atcccccgag ttgggt	. 36
Egggegeaca Eggeogggeo accessory	
<210> 95	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 95	
cgcgcgcaca tgtctatatc ccccgagttg ggtccc	36
egegegeara tyterature ceeegages syvere	
<210> 96	
<211> 36	
<212> DNA	
<213> Synthetic	
4400- 96	
<400> 96	36
cgcgcgcaca tgtcttcccc cgagttgggt cccacc	
<210> 97	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 97	
	36
cgcgcgcaca tgtctcccga gttgggtccc accttg	
<210> 98	
<211> 36	
<212> DNA	
<213> Synthetic	
. ===	
<400> 98	
cgcgcgcaca tgtctgagtt gggtcccacc ttggac	36
egegegeaca egecegagee gggeteened tegg-	
<210> 99	
<211> 36	
<212> DNA	
<213> Synthetic	•
<400> 99	
cgcgcgcaca tgtctttggg tcccaccttg gacaca	3
egegegeaca egeceeeggg receased	,
100	
<210> 100	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 100	
cgcgcgcaca igictggicc caccitggac acactg	3
- chelebrack -diseasant canaritans assess	_
MA 6- 101	
<210> 101	
<211> 36	
<212> DNA	
<213> Synthetic	
4400-15	

egegegeaca tgteteccae ettggacaca etgcag		36
<210> 102		
<211> 36		
<212> DNA		•
<213> Synthetic		
400- 102		
<400> 102 cgcgcgcaca tgtctacctt ggacacactg cagctg		36
	÷	
<210> 103	•	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 103	*	
cgcgcgcaca tgtctttgga cacactgcag ctggac		36
<210> 104		
<211> 36		
<212> DNA	•	
<213> Synthetic	•	
•		
<400> 104		36
cgcgcgcaca tgtctgacac actgcagctg gacgtc		
<210> 105	•	•
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 105		
cgcgcgcaca tgtctacact gcagctggac gtcgcc		36
<210> 106	·	
<211> 36		
<212> DNA <213> Synthetic		
(213) Synchectic		
<400> 106		
cgcgcgcaca tgtctctgca gctggacgtc gccgac		36
-2105 107	· .	
<210> 107 . <211> 36		
<212> DNA		
<213> Synthetic		
<400> 107 cgcgcgcaca tgtctcagct ggacgtcgcc gacttt	•	36
cgcgcgcaca tgtcttdgtt gg-og-og-o		
<210> 108		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 108	•	
cgcgcgcaca tgtctctgga cgtcgccgac tttgcc		36
<210> 109		
<211> 36		
<212> DNA		
<213> Synthetic		
	•	
<400> 109		36
cgcgcgcaca tgtctgacgt cgccgacttt gccacc		
<210> 113	• •	
<211> 36		
<212> DNA		

<213> Synthetic	
<400> 110	
cgcgcgcaca tgtctgtcgc cgactttgcc accacc	36
<210> 111	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 111	
cgcgcgcaca tgtctgccga ctttgccacc accatc	36
- cgcgcgcaca - cgccgcaca - cgcgcgcaca	
<210> 112	* .
<211> 36	
<212> DNA	
<213> Synthetic	
· · · · · · · · · · · · · · · · · · ·	•
<400> 112	3.5
cgcgcgcaca tgtctgactt tgccaccacc atctgg	36
<210> 113	•
<211> 36	
<212> DNA	
<213> Synthetic	177
<400> 113	
cgcgcgcaca tgtcttttgc caccaccatc tggcag	36
egegegeded egeteetge terrerer in a	
<210> 114	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 114	36
cgcgcgcaca tgtctgccac caccatctgg cagcag	30
.010- 115	
<210> 115 <211> 36	
<211> 30 <212> DNA	
<213> Synthetic	
<400> 115	
cgcgcgcaca tgtctaccac catctggcag cagatg	. 36
	·
<210> 116	
<211> '36	
<212> DNA	·
<213> Synthetic	
100-116	
<400> 116 cgcgcgcaca tgtctaccat ctggcagcag atggaa	. 36
cgcgcgcaca tytttactat toggcagtag atggaa	
<210> 117	
<211> 36	
<212> DNA	
<213> Synthetic	•
-	•
<400> 117	
cgcgcgcaca tgtctatctg gcagcagatg gaagaa	36
<210> 118	
<211> 36	
<212> DNA	•
<213> Synthetic	
400x 110	
<400> 118	.31
cgcgcgraca tgtcttggca gcagatggaa gaactg	

<210> 119				
<211> 36				
<212> DNA				
<213> Synthetic				
				•
<400> 119				36
gegegeaca tgteteagea gatggaagaa etggga				
<210> 120				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 120 cgcgcgcaca tgtctcagat ggaagaactg ggaatg				36
egegegeaca egectonger panagar p				
<210> 121				
<211> 36				
<212> DNA				
<213> Synthetic				
•				
<400> 121			•	3.6
cgcgcgcaca tgtctatgga agaactggga atggcc		:		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
<210> 122	•			
<211> 36				
<212> DNA				
<213> Synthetic	•			
(213) Byllelico20			·	
<400> 122				36
cgcgcgcaca tgtctgaaga actgggaatg gcccct	<u>:</u>			20
<210> 123			٠.	
<211> 36				
<212> DNA			•	
<213> Synthetic				٠.
1213 Synthetic				
<400> 123		•		. 36
cgcgcgcaca tgtctgaact gggaatggcc cctgc	<b>5</b>		•	30
<210> 124				
<211> 36	* *			
<211> 36 <212> DNA		• 0		
<213> Synthetic				
<400> 124		**		
cgcgcgcaca tgtctctggg aatggcccct gccct	g			36
<210> 125				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 125				
cgcgcgcaca tgtctggaat ggcccctgcc ctgca	ıg .			.36
cgogogocou oposogo				
<210> 126		•		
<211> 36				
<212> DNA				
<213> Synthetic		•		
<400> 126	cc			36
egegegeaca tgtetatgge ecetgecetg cage				
<210> 127		-		
<211> 36				
<212> DNA				

<213> Synthetic

cgcgcgcaca tgtctgcccc tgccctgcag cccacc	36
<210> 128	
<211> 36	
<212> DNA	
<213> Synthetic	
3,200	
<400> 128	
cgcgcgcaca tgtctcctgc cctgcagccc acccag	36
egegegeata tyteteetye eeryeesoo bessel	
220- 120	
<210> 129	•
<211> 36	
<212> DNA	
<213> Synthetic	
120	
<400> 129	3.6
cgcgcgcaca tgtctgccct gcagcccacc cagggt	
<210> 130	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 130	36
cgcgcgcaca tgtctctgca gcccacccag ggtgcc	30
·	
<210> 131	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 131	26
egegegeaca tgteteagee cacceagggt gecatg	36
•	
<210> 132	
<211> 36	
<212> DNA	
<212> DNA	
<212> DNA <213> Synthetic <400> 132	36
<212> DNA <213> Synthetic	36
<212> DNA <213> Synthetic <400> 132	36
<212> DNA <213> Synthetic <400> 132	36
<212> DNA <213> Synthetic  <400> 132  cgcgcgcaca tgtctcccac ccagggtgcc atgccg	36
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133	36
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36	36
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA	36
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA	
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA <213> Synthetic  <400> 133	36
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA <213> Synthetic	
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA <213> Synthetic  <400> 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc	
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA <213> Synthetic  <400> 133	
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA <213> Synthetic  <400> 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  <210> 134 <211> 36	
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA <213> Synthetic  <400> 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  <210> 134 <211> 36 <212> DNA <211> 36 <212> DNA <213> Synthetic	
<212> DNA <213> Synthetic  <400> 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  <210> 133 <211> 36 <212> DNA <213> Synthetic  <400> 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  <210> 134 <211> 36	
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	36
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	36
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;211&gt; 36 &lt;211&gt; 36 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	36
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;211&gt; 37 &lt;211&gt; 36 &lt;211&gt; 36 &lt;211&gt; 36 &lt;211&gt; 30 &lt;211&gt; 31 &lt;211&gt; 30 &lt;211&gt; 31 &lt;211&gt; 31<!--21--> 31 &lt;211&gt; 32 &lt;211&gt; 31<!--21--> 32 &lt;211&gt; 32 &lt;211&gt; 33 &lt;211&gt; 34 &lt;211&gt; 35<!--21--> 34 &lt;211&gt; 35<!--21--> 36<!--21--> 31<!--2--> 31<!--2--> 31</pre>	36
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400 134 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 134 cgcgcgcaca tgtctcaggg tgccatgccg gccttc  &lt;210&gt; 135 &lt;211&gt; 36</pre>	36
<pre> &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 134 cgcgcgcaca tgtctaggg tgccatgccg gccttc  &lt;210&gt; 135 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	36
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400 134 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 134 cgcgcgcaca tgtctcaggg tgccatgccg gccttc  &lt;210&gt; 135 &lt;211&gt; 36</pre>	36
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 134 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 134 cgcgcgcaca tgtctcaggg tgccatgccg gccttc  &lt;210&gt; 135 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;300&gt; Synthetic &lt;210&gt; 135 &lt;211&gt; 35 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;210&gt; 135 &lt;211&gt; 35 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	36
<pre>&lt;212&gt; DNA</pre>	36
<pre>&lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 132 cgcgcgcaca tgtctcccac ccagggtgcc atgccg  &lt;210&gt; 133 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 133 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 134 cgcgcgcaca tgtctaccca gggtgccatg ccggcc  &lt;210&gt; 134 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 134 cgcgcgcaca tgtctcaggg tgccatgccg gccttc  &lt;210&gt; 135 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;300&gt; Synthetic &lt;210&gt; 135 &lt;211&gt; 35 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;210&gt; 135 &lt;211&gt; 35 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	36
<pre>&lt;212&gt; DNA</pre>	36

<212> DNA		
<213> Synthetic		
	·	
<400> 136		
cgcgcgcaca tgtctgccat gccggccttc gcctct		36
cgcgcgcaca cgcccgccac geoggeness general		•
<210> 137	•	
<211> 36		
<212> DNA	·	
<213> Synthetic		
•		
<400> 137		3.6
cgcgcgcaca tgtctatgcc ggccttcgcc tctgct		36
<210> 138	101	
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 138		
cgcgcgcaca tgtctccggc cttcgcctct gctttc		36
<210> 139	•	
<211> 36		
<212> DNA	-	
<213> Synthetic	•	
-0-24		
<400> 139		
cgcgcgcaca tgtctgcctt cgcctctgct ttccag		36
egegegeaca egeegeee egeeses		
<210> 140		
<211> 36	·	
<212> DNA		
<213> Synthetic	•	
-100> 140	• •	
<400> 140	•	36
cgcgcgcaca tgtctttcgc ctctgctttc cagcgc	· ·	
<210> 141	•	
<211> 36		
<212> DNA		
<213> Synthetic		
•		
<400> 141	•	36
cgcgcgcaca tgtctgcctc tgctttccag cgccgg	•	50
	· ·	
<210> 142		
<211> 36		
<212> DNA	•	
<213> Synthetic		
·	•	
<400> 142	•	2.0
cgcgcgcaca tgtcttctgc tttccagcgc cgggca		36
<210> 143	*	
<211> 36		
<212> DNA		-
<213> Synthetic		
<400> 143	•	
cgcgcgcaca tgtctgcttt ccagcgccgg gcagga		36
cholodowo chorodores contacting the		
<210> 144		
<210> 144		
<211> 36	•	
<212> DNA	•	
<213> Synthetic	*	
<400> 144		36
cocococaca tototttoca gogoogggca ggaggg		٥

	<210> 145					
	<211> 36					
	<212> DNA	3.0				
	<213> Synthetic					
	<213> Synthetic					
	<400> 145					
cgcgc	gcaca tgtctcagcg	ccgggcagga	ggggtc			36
			•			
	<210> 146	•				
	<211> 36					
	<212> DNA					
				•		
•	<213> Synthetic				•	
·				·		
	<400> 146					
cgcgc	gcaca tgtctcgccg	ggcaggaggg	gtcctg			36
		•				
	<210> 147					
	<211> 36					
	<212> DNA					
	<213> Synthetic	•				
	12132 Synchictic		•	·		
	.400- 147					
•	<400> 147					3.0
cgcgc	gcaca tgtctcgggc	aggaggggtc	ctggtt	•		36
	<210> 148	•			•	
	<211> 36					
	<212> DNA					
	<213> Synthetic					
	<400> 148			•		
		2000010010	artact			36
egege	gcaca tgtctgcagg	aggggccccg	gregee			
			٠.			
	<210> 149					
	<211> 36					
	<212> DNA					
	<213> Synthetic		-			
	<400> 149				•	
cgcge	cgcaca tgtctggagg	ggtcctggtt	gctagc	•	•	36
					•	
	<210> 150		•			
	<211> 36				* * * * *	
	<212> DNA					
				•		
	<213> Synthetic	i		•		
					•	
	<400> 150		1 Page 1			
cgcg	cgcaca tgtctggggt	cctggttgct	agccat			36
	<210> 151			· · · · · · · · · · · · · · · · · · ·		
	<211> 36					
	<212> DNA	*	•			•
	<213> Synthetic	:				
	<400> 151					-
	cgcaca tgtctgtcct	. aartacteac	Catcto			36
cgcg	cycaca cycciyccc	. ggccgcccgc			•	
	.010- 150					
	<210> 152					
	<211> 36		•			
	<212> DNA	•				
	<213> Synthetic	:				
				. *		
	<400> 152					
coce	cgcaca tgtctctggt	tgctagccat	ctgcag			3.6
-9-9	-2-0-0 -3-0-0-033					
	<210> 153					
		•				
	<211> 36					
	<212> DNA			•		
	-2125 Simphoria					

<400> 153			
cgcgcgcaca tgtctgttgc tagccatc	tg cagage		36
			•
<210> 154			,
<211> 36			
<212> DNA		•	
<213> Synthetic			
<213> Synchectic			
400- 154	**		
<400> 154		•	36
cgcgcgcaca tgtctgctag ccatctgc	ay ayerer		
200 455			•
<210> 155		• •	,
<211> 36 /			•
<212> DNA		•	
<213> Synthetic			
<400> 155			36
cgcgcgcaca tgtctagcca tctgcaga	age tteetg		
	.•		
<210> 156			
<211> 36			,
<212> DNA		· :	
<213> Synthetic			
•			
<400> 156	•		2.5
cgcgcgcaca tgtctcatct gcagage	ttc ctggag	•	. 36
		•	
<210> 157			
<211> 36			
<212> DNA			
<213> Synthetic			
-100- 157			
<400> 157	cta' ascata		36
cgcgcgcaca tgtctctgca gagcttc	cra araaca	4	
<210> 158			
<211> 36			
<211> 50 <212> DNA	•		
<213> Synthetic		·	
4225° 53			
<400> 158	•	. *	
cgcgcgcaca tgtctcagag cttcctg	gag gtgtcg		36
<210> 159			
<211> 36	i		•
<212> DNA			
<213> Synthetic		•	
		•	•
<400> 159	101	•	
cgcgcgcaca tgtctagctt cctggag	gtg tcgtac	* * .	36
<210> 160 ·			
<211> 36		•	
<212> DNA		,	•
<213> Synthetic	• •		* .
<400> 160			
egegegeaca tgtettteet ggaggtg	tcg taccgc		36
			•
<210> 161			
<211> 36			
<212> DNA		•	
<213> Synthetic		•	
•			
<400> 161			
cgcgcgcaca tgtctctgga ggtgtcg	tac -cgcgtt		36
212. 152			

<211> 36		
<212> DNA		
<213> Synthetic		
•		
<400> 162		
cgcgcgcaca tgtctgaggt gtcgtaccgc gttcta		36
cycycycaca cycocyayya yaayaa a		
<210> 163 .		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 163	•	30
cgcgcgcaca tgtctgtgtc gtaccgcgtt ctacgc		36
<210> 164		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 164	•	
cgcgcgcaca tgtcttcgta ccgcgttcta cgccac		36
- cycycy cuca type to the cycycy		
<210> 165	.*	
<211> 36		
<212> DNA		
<213> Synthetic	•	
	•	
<400> 165		36
cgcgcgcaca tgtcttaccg cgttctacgc cacett		٠,
<210> 166		
<211> 36		
<212> DNA	•	
<213> Synthetic		
<400> 166		_
cgcgcgcaca tgtctcgcgt tctacgccac cttgcg	•	3 (
	• •	•
<210> 167	•	
<211> 36		
<212> DNA		
<213> Synthetic		
	•	
<400> 167		
cgcgcgcaca tgtctgttct acgccacctt gcgcag	•	3
cgagagaaa agaaagaaa	* *	
<210> 158		
<211> 36	•	
<211> DNA		
<213> Synthetic	•	
-400- 100		
<400> 168	٠.	3
cgcgcgcaca tgtctctacg ccaccttgcg cagccc		
	•	
<210> 169		
<211> 36	•	
<212> DNA		
<213> Synthetic	,	
in the second se		
<400> 169		
cgcgcgcaca tgtctcgcca ccttgcgcag cccgac	•	3
<210> 170		
<211> 36		
<211> 30 <212> DNA	•	
<212> DNA <213> Synthetic	•	
<513> SAUCHECTC		
<400> 170		
<400> 1/0		

cgcgcgcaca tgtctcacct	tgcgcagccc	gacatg		36
<210> 171				
<211> 36				
<211> 30 <212> DNA	•			
<213> Synthetic				
	•			
<400> 171				
cgcgcgcaca tgtctcttgc	gcagcccgac	atggct		36
• •				
<210> 172				
<211> 36				
<212> DNA				
<213> Synthetic				
<400> 172				
cgcgcgcaca tgtctgcgca	gecegacatg	gctaca		36
cgcgcgcau- cgcccgcgc			·	
<210> 173			.*	
<211> 36				
<212> DNA		•		
<213> Synthetic				
			•	
<400> 173	casastance	20200		36
cgcgcgcaca tgtctcagcc	egacatgget	acacca	•	
<210> 174		•		
<211> 36				
<212> DNA	•			
<213> Synthetic			•	
•				
<400> 174				36
cgcgcgcaca tgtctcccga	catggctaca	ccattg		50
<210> 175				
<211> 36		•		
<212> DNA				
<213> Synthetic		•	. •	
<400> 175				3.0
cgcgcgcaca tgtctcgcca	ccttgcgcag	cccact		36
.010- 126				
.<210> 176 <211> 36			·	
<211> 30 <212> DNA			•	
<213> Synthetic			•	
<400> 176			·	
cgcgcgcaca tgtctcacct	tgegeagee	actagt	•	36
-216- 177			•	
<210> 177 <211> 36	•			
<211> 36 <212> CMA				
<213> Synthetic	:			
•				
<400> 177				
cgcgcgcaca tgtctcttgc	gcagcccact	agtcat		36
<210> 178				
<211> 36 <212> DNA				
<213> Synthetic				
	=			
<400> 178				
cgcgcgcaca tgtctgcgca	geceaetag	cateca		3 €
• • • • • • • • • • • • • • • • • • • •				
<210> 173	•	•		
<211> 36			•	
<212> DNA			•	

<213> Synthetic		
<400> 179		
cgcgcgcaca tgtctcagcc cactagtcat ccacct.		36
cgcgcgca		
<210> 180		
<21i> 36		
<212> DNA		
<213> Synthetic		
<400> 180		36
cgcgcgcaca tgtctcccac tagtcatcca cctatg		30
<210> 181		
<211> 36 <212> DNA		
<213> Synthetic		
(21) Dynamour		
<400> 181		
cgcgcgcaca tgtctcgcca ccttgcgcag cccggc		3.6
<210> 182	•	
<211> 36	•	
<212> DNA	•	
<213> Synthetic		
<400> 182		
cgcgcgcaca tgtctcacct tgcgcagccc ggcggc		36
cgcgcgcaca agacocco cgcgcgcaca		
<210> 183		
<211> 36		
<212> DNA	•	
<213> Synthetic		
.400- 303		
<400> 183 cgcgcgcaca tgtctcttgc gcagcccggc ggcggc		3.6
Cacacacaca caccacaca acadocada assessa		
<210> 184		
<211> 36		
<212> DNA		
<213> Synthetic		-
<400> 184		3.6
cgcgcgcaca tgtctgcgca gcccggcggc ggctct		
<210> 185		
<211> 36		
<212> DNA		
<213> Synthetic		
	•	
<400> 185		3.6
egegegeaca tgteteagee eggeggegge tetgae		
<210> 186		
<211> 36		
<212> DNA	•	
<213> Symphesic		•
<\$?₹> 186		20
cgcgcgcaca tgtctcccgg cggcggctct gacatg		36
.010. 107		•
<210> 187		
<211> 35 <212> DNA		
<pre>&lt;212&gt; DMA &lt;213&gt; Symthetic</pre>		
- man - mil representation		
<400> 187		
tatatatece geogrageca tetoggecte cecae		3:6

<210> 188	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 188	
tatatatgcg gccgcagcca tgtcacgcgt acgatt	36
<210> 189	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 189	
tatatatgcg gccgcagcca tgtcagagcc gccgcc	36
Catatacycy googouston opening to the	
<210> 190	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 190	36
tatatatgcg gccgctgtag ccatgtcggg ctgcgc	
-2105 191	
<210> 191 <211> 36	
<211> 50 <212> DNA	
<213> Synthetic	
<400> 191	7.6
tatatatgcg gccgctgtag ccatgtcacg cgtacg	36
<210> 192	
<211> 36	
<212> DNA	
<213> Synthetic	
<213> Synthetic	
<213> Synthetic <400> 192	36
<213> Synthetic	36
<213> Synthetic <400> 192	36
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36	36
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA	3'6
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36	36
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA <213> Synthetic	36
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193  <211> 36  <212> DNA  <213> Synthetic  <400> 193	36
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA <213> Synthetic	
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA <213> Synthetic  <400> 193 tatatatgcg gccgctggtg tagccatgtc gggctg	
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193  <211> 36  <212> DNA  <213> Synthetic  <400> 193	
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA <213> Synthetic  <400> 193 tatatatgcg gccgctggtg tagccatgtc gggctg  <210> 194	
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA <213> Synthetic  <400> 193 tatatatgcg gccgctggtg tagccatgtc gggctg  <210> 194 <211> 36	
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA <213> Synthetic  <400> 193 tatatatgcg gccgctggtg tagccatgtc gggctg  <210> 194 <211> 36 <212> DNA <211> 36 <212> DNA	
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA <213> Synthetic  <400> 193 tatatatgcg gccgctggtg tagccatgtc gggctg  <210> 194 <211> 36 <212> DNA <213> Synthetic  <400> 194 <211> 36 <212> DNA <213> Synthetic  <400> 194	36
<213> Synthetic  <400> 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  <210> 193 <211> 36 <212> DNA <213> Synthetic  <400> 193 tatatatgcg gccgctggtg tagccatgtc gggctg  <210> 194 <211> 36 <212> DNA <213> Synthetic	
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 193 tatatatgcg gccgctggtg tagccatgtc gggctg  &lt;210&gt; 194 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 194 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 194 &lt;211&gt; 36</pre> <pre>&lt;210&gt; degree of tagccatgtc gggctg</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 193 tatatatgcg gccgctggtg tagccatgtc gggctg  &lt;210&gt; 194 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic  &lt;400&gt; 194 tatatatgcg gccgctggtg tagccatgtc acgcgt  &lt;400&gt; 194 tatatatgcg gccgctggtg tagccatgtc acgcgt  &lt;210&gt; 195 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 195 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36
<pre>&lt;213&gt; Synthetic  &lt;400&gt; 192 tatatatgcg gccgctgtag ccatgtcaga gccgcc  &lt;210&gt; 193</pre>	36

<400> 196	
tatatatgcg gccgccaatg gtgtagccat gtcggg	36
<210> 197	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 197	36
tatatatgcg gccgccaatg gtgtagccat gtcacg	30
•	
<210> 198	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 198	
tatatatgcg gccgccaatg gtgtagccat gtcaga	36
<210> 199	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 199	
tatatatgcg gccgcgcca atggtgtagc catgtc	3.6
tatatatycy geogeocea acygogouse constraint	
<210> 200	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 200	36
tatatatgcg gccgcagggc ccaatggtgt agccat	
.010. 201	
<210> 201	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 201	36
tatatatgeg geegeggeag ggeecaatgg tgtage	
<210> 202	
<211> 36	
<212> DNA	
<213> Synthetic	
<400> 202	
tatatatgeg geegegetgg cagggeecaa tggtgt	3.6
<210> 203	
<211> 36	,
<212> DNA	
<2:3> Synthetic	
438> 203	
tatatatgcg gccgcggagc tggcagggcc caatgg	36
<210> 204	•
<211> 3£	
<212> DNA	
<213> Synthetic	
-nen- niwamana	
<400> 204	
	3.6
tatatatgog geogecaggg Agottgatagg geocaa	
411.0× 20E	
<210> 205 <211> 36	
**     P   ND	

<212> DNA					
<213> Synthetic					
(213) Synthetic				•	
<400> 205					36
tatatatgcg gccgcgggca	gggagctggc	agggcc			96
<210> 206					
<211> 36					
<212> DNA					
<213> Synthetic					
<400> 206					
		######################################			36
tatatatgcg gccgcctggg	geagggagee	ggcagg			•
<210> 207					
<211> 36					
<212> DNA					
<213> Synthetic					
•					
<400> 207	•				
tatatatgcg gccgcgctct	ggggcaggga	getgge			36
tatatatgeg geograpite	3533233-	5 .55			
-210- 208					
<210> 208				•	
<211> 36					
<212> DNA					•
<213> Synthetic	•				
				•	
<400> 208					
tatatatgcg gccgcgaagc	tctggggcag	ggagct			36
<210> 209					
<211> 36					
<212> DNA					
<213> Synthetic					
<\$13> Synthetic	•				
• •					
<400> 209					36
		caggga			36
<400> 209		caggga			36
<400> 209		caggga			36
<400> 209 tatatatgcg gccgccagga		caggga			36
<400> 209 tatatatgcg gccgccagga		caggga			36
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36 <212> DNA	a agetetgggg	caggga		·	36
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36	a agetetgggg	caggga		·	36
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36 <212> DNA <213> Synthetic	a agetetgggg	caggga		·	36
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36 <212> DNA <213> Synthetic	a agetetgggg				36
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36 <212> DNA <213> Synthetic	a agetetgggg				
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36 <212> DNA <213> Synthetic <400> 210 tatatatgcg gccgcgagca	a agetetgggg				
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36 <212> DNA <213> Synthetic <400> 210 tatatatgcg gccgcgagca <210> 211	a agetetgggg				
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36 <212> DNA <213> Synthetic <400> 210 tatatatgcg gccgcgagca <210> 211 <211> 36	a agetetgggg				
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga &lt;210&gt; 210</pre>	a agctctgggg				
<400> 209 tatatatgcg gccgccagga <210> 210 <211> 36 <212> DNA <213> Synthetic <400> 210 tatatatgcg gccgcgagca <210> 211 <211> 36	a agctctgggg				
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga &lt;210&gt; 210</pre>	a agctctgggg				
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga &lt;210&gt; 210</pre>	a agctctgggg				36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga &lt;210&gt; 210</pre>	a agetetgggg	1 gggcag			
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga &lt;210&gt; 210</pre>	a agetetgggg	1 gggcag			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	1 gggcag			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	1 gggcag			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	1 gggcag			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	1 gggcag	*		36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	1 gggcag	*		36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	1 gggcag			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agctctgggg	crââââ	*		36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agctctgggg	crââââ			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agctctgggg	crââââ			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agctctgggg	crââââ			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga &lt;210&gt; 210</pre>	a agctctgggg	crââââ			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga &lt;210&gt; 210</pre>	a agctctgggg	crââââ			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	crââââ			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	crââââ			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga &lt;210&gt; 210</pre>	a agetetgggg	crââââ			36
<pre>&lt;400&gt; 209 tatatatgcg gccgccagga  &lt;210&gt; 210</pre>	a agetetgggg	a getetg			36

<210> 214			
<211> 36			
<212> DNA			
<213> Synthetic			
-			
<400> 214			36
tatatatgeg geegeeteta aagaettgag caggaa			30
<210> 215			
<211> 36			
<212> DNA			
<213> Synthetic			
<400> 215			
tatatatgcg gccgcttgct ctaaagactt gagcag			36
<210> 216			
<211> 36			
<212> DNA			
<213> Synthetic			
<400> 216			•
tatatatgcg gccgccactt gctctaaaga cttgag			36
Cacacacage good on the cacacacage good on the cacacacage good on the cacacacacacacacacacacacacacacacacacaca			
<210> 217			
<211> 36			
<212> DNA			
<213> Synthetic			
<400> 217			
			36
tatatatgeg geegeeetea ettgetetaa agaett			
<210> 218			
<211> 36			
<212> DNA			
<213> Synthetic			
<400> 218			36
tatatatgcg gccgccttcc tcacttgctc taaaga			30
<210> 219			
<211> 36			
<212> DNA <213> Synthetic			
(21) Synchesis			
<400> 219			
tatatatgcg gccgcgatct tcctcacttg ctctaa			36
.010- 220			
<210> 220 <211> 36			
<212> DNA <213> Synthetic			
Test by			
<400> 220			
tatatatgcg gccgcctgga tcttcctcac ttgctc			36
.7* 7 221			
<210> 221		•	
<211> 36			
<212> DNA			
<213> Synthetic			
<400> 221			
tatatatgcg gccgcgccct ggatcttcct cacttg			3
<213> 222		•	
<211> 36	*		
<212> DNA			

	•	
<400> 222	•	
tatatatgcg gccgcatcgc	cctggatctt cctcac	36
<210> 223		
<211> 36		
<212> DNA		
<213> Synthetic		
	• '	
<400> 223		
tatatatgcg gccgcgccat	egecetggat etteet	36
<210> 224		
<211> 36		
<212> DNA		
<213> Synthetic		
-4005 224		
<400> 224	17.1	26
tatatatgcg gccgctgcgc	categeetg gatett	36
	* • • • • • • • • • • • • • • • • • • •	
<210> 225		
<211> 36		
	1	
<212> DNA	•	
<213> Synthetic		
<400> 225		
	encestence etanst	36
tatatatgcg gccgccgctg	Cyccatoget traggat	
<210> 226		
<211> 36		
<212> DNA		
<213> Synthetic		
<400> 226		
tatatatgcg gccgcgagcg	ctococcatc occcto	36
ratatating accaesance	ougestant series	
<210> 227	•	
<211> 36		
<212> DNA		
<213> Synthetic	• "	
(215) Synchectic		
	•	
<400> 227		
tatatatgcg gccgcctgga	gegetgegee ategee	36
<210> 228		
<211> 36		
<212> DNA		
<213> Synthetic		
<213> Synthetic		
<213> Synthetic		3:6
<213> Synthetic		3.6
<213> Synthetic		36
<213> Synthetic		36
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229		36
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229 <211> 36		36
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229 <211> 36 <212> DNA	ggagcgctgc gccatc	36
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229 <211> 36	ggagcgctgc gccatc	36
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229 <211> 36 <212> DNA	ggagcgctgc gccatc	36
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229 <211> 36 <212> DNA <213> Synthetic	ggagcgctgc gccatc	36
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229 <211> 36 <212> DNA <213> Synthetic <400> 229	ggagegetge gecate	
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229 <211> 36 <212> DNA <213> Synthetic	ggagegetge gecate	36
<213> Synthetic  <400> 228 tatatatgcg gccgcctcct  <210> 229  <211> 36  <212> DNA  <213> Synthetic  <400> 229 tatatatgcg gccgccttct	ggagegetge gecate	
<213> Synthetic <400> 228 tatatatgcg gccgcctcct <210> 229 <211> 36 <212> DNA <213> Synthetic <400> 229	ggagegetge gecate	
<213> Synthetic  <400> 228 tatatatgcg gccgcctcct  <210> 229  <211> 36  <212> DNA  <213> Synthetic  <400> 229 tatatatgcg gccgccttct	ggagegetge gecate	
<pre>&lt;213&gt; Synthetic &lt;400&gt; 228 tatatatgcg gccgcctcct &lt;210&gt; 229 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 229 tatatatgcg gccgccttct &lt;210&gt; 230 &lt;211&gt; 36</pre>	ggagegetge gecate	
<pre>&lt;213&gt; Synthetic &lt;400&gt; 228 tatatatgcg gccgcctcct &lt;210&gt; 229 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 229 tatatatgcg gccgcctct &lt;210&gt; 230 &lt;211&gt; 36 &lt;212&gt; DNA</pre>	t cetggagege tgegee	
<pre>&lt;213&gt; Synthetic &lt;400&gt; 228 tatatatgcg gccgcctcct &lt;210&gt; 229 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 229 tatatatgcg gccgccttct &lt;210&gt; 230 &lt;211&gt; 36</pre>	t cetggagege tgegee	
<pre>&lt;213&gt; Synthetic &lt;400&gt; 228 tatatatgcg gccgcctcct &lt;210&gt; 229 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 229 tatatatgcg gccgccttct &lt;210&gt; 230 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	t cetggagege tgegee	
<pre>&lt;213&gt; Synthetic &lt;400&gt; 228 tatatatgcg gccgcctcct &lt;210&gt; 229 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 229 tatatatgcg gccgcctct &lt;210&gt; 230 &lt;211&gt; 36 &lt;212&gt; DNA</pre>	t cetggagege tgegee	
<pre>&lt;213&gt; Synthetic &lt;400&gt; 228 tatatatgcg gccgcctcct &lt;210&gt; 229 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic &lt;400&gt; 229 tatatatgcg gccgccttct &lt;210&gt; 230 &lt;211&gt; 36 &lt;212&gt; DNA &lt;213&gt; Synthetic</pre>	ggagegetge gecate  t cetggagege tgegee	

<210> 231

## This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:
BLACK BORDERS
$\square$ image cut off at top, bottom or sides
FADED TEXT OR DRAWING
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
$\square$ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
Полить

## IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.